

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 15, 2006, 03:54:31 ; Search time 243 Seconds

(without alignments)

3444.590 Million cell updates/sec

Title: US-10-018-105A-3

Perfect score: 3523

Sequence: 1 tataataatataatata.....tcaattatccatataatgc 3525

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 290227 seqs, 117694381 residues

Total number of hits satisfying chosen parameters: 580454

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_NA_New:*

1: /EMC_Caliera_SIDS3/ptodata/1/pubpna/us09_NEW_pub.seq:*

2: /EMC_Caliera_SIDS3/ptodata/1/pubpna/us06_NEW_pub.seq:*

3: /EMC_Caliera_SIDS3/ptodata/1/pubpna/us07_NEW_pub.seq:*

4: /EMC_Caliera_SIDS3/ptodata/1/pubpna/us08_NEW_pub.seq:*

5: /EMC_Caliera_SIDS3/ptodata/1/pubpna/PCM_NEW_pub.seq:*

6: /EMC_Caliera_SIDS3/ptodata/1/pubpna/us11_NEW_pub.seq:*

7: /EMC_Caliera_SIDS3/ptodata/1/pubpna/us01_NEW_pub.seq:*

8: /EMC_Caliera_SIDS3/ptodata/1/pubpna/us00_NEW_pub.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description | |
|------------|-------|-------------|--------|------------------------|----------------------|-------------------|
| 1 | 408.8 | 11.6 | 2571 | 7 US-11-217-529-5831 | Sequence 5831, APP | |
| 2 | 403.8 | 11.5 | 2574 | 7 US-11-217-529-78462 | Sequence 2281, APP | |
| 3 | 396 | 11.2 | 2580 | 7 US-11-217-529-78462 | Sequence 78462, APP | |
| 4 | 383 | 10.9 | 2577 | 7 US-11-217-529-78424 | Sequence 78424, APP | |
| 5 | 127.4 | 3.6 | 2910 | 7 US-11-217-529-7745 | Sequence 7745, APP | |
| 6 | 116.2 | 3.3 | 2910 | 7 US-11-217-529-7745 | Sequence 3077, APP | |
| 7 | 82.2 | 2.3 | 1134 | 7 US-11-217-529-166160 | Sequence 166160, APP | |
| 8 | 79.8 | 2.3 | 9839 | 6 US-10-517-441-456 | Sequence 456, APP | |
| 9 | 79.8 | 2.2 | 9839 | 6 US-10-517-441-730 | Sequence 710, APP | |
| c | 10 | 78.2 | 2.2 | 8759 | 6 US-10-517-441-256 | Sequence 256, APP |
| c | 11 | 78.2 | 2.2 | 8759 | 6 US-10-517-441-530 | Sequence 530, APP |
| c | 12 | 70.2 | 2.0 | 5286 | 6 US-10-517-441-234 | Sequence 234, APP |
| c | 13 | 70.2 | 2.0 | 5286 | 6 US-10-517-441-456 | Sequence 568, APP |
| c | 14 | 1.9 | 19634 | 6 US-10-517-441-688 | Sequence 688, APP | |
| c | 15 | 65.6 | 1.8 | 2388 | 6 US-10-517-441-370 | Sequence 370, APP |
| c | 16 | 63.6 | 1.8 | 2388 | 6 US-10-517-441-664 | Sequence 644, APP |
| c | 17 | 63.6 | 1.8 | 15355 | 6 US-10-517-441-415 | Sequence 415, APP |
| c | 18 | 63.6 | 1.8 | 15355 | 6 US-10-517-441-659 | Sequence 659, APP |
| c | 19 | 62.6 | 1.8 | 6432 | 6 US-10-517-441-605 | Sequence 605, APP |
| c | 20 | 62.4 | 1.8 | 12610 | 6 US-10-517-441-411 | Sequence 411, APP |
| c | 21 | 62.4 | 1.8 | 12610 | 6 US-10-517-441-715 | Sequence 715, APP |
| c | 22 | 62.2 | 1.8 | 6343 | 6 US-10-517-441-666 | Sequence 616, APP |
| c | 23 | 61.4 | 1.7 | 10865 | 6 US-10-517-441-464 | Sequence 464, APP |
| c | 24 | 61.4 | 1.7 | 10865 | 6 US-10-517-441-738 | Sequence 738, APP |
| c | 25 | 61 | 1.7 | 5286 | 6 US-10-517-441-234 | Sequence 234, APP |
| c | 26 | 61 | 1.7 | 19634 | 6 US-10-517-441-568 | Sequence 568, APP |
| c | 27 | 60.8 | 1.7 | 19634 | 6 US-10-517-441-414 | Sequence 614, APP |
| c | 28 | 60 | 1.7 | 12610 | 6 US-10-517-441-716 | Sequence 442, APP |
| c | 29 | 60 | 1.7 | 12610 | 6 US-10-517-441-501 | Sequence 716, APP |
| c | 30 | 60 | 1.7 | 4001 | 6 US-10-517-441-501 | Sequence 501, APP |

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Gapext 1.0

ALIGNMENTS

Copyright (c) 1993 - 2006 Biocceleration Ltd.

Om nucleic - nucleic search, using sw model
Run on: June 15, 2006, 03:11:16 : Search time 653 Seconds
(without alignments)
10100.59 Million cell updates/sec

Title: US-10-018-105A-3
Perfect score: 3523
Sequence: 1 tataattataatataataataataatattgc 3525
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched:

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing Minimum Match 0%

Listing first 45 summaries

Database :

- 1: /EMC_Celerra_SIDS3/podata/2/ina/1_COMB.seq:*
- 2: /EMC_Celerra_SIDS3/podata/2/ina/5A_COMB.seq:*
- 3: /EMC_Celerra_SIDS3/podata/2/ina/6A_COMB.seq:*
- 4: /EMC_Celerra_SIDS3/podata/2/ina/6B_COMB.seq:*
- 5: /EMC_Celerra_SIDS3/podata/2/ina/7_COMB.seq:*
- 6: /EMC_Celerra_SIDS3/podata/2/ina/H_COMB.seq:*
- 7: /EMC_Celerra_SIDS3/podata/2/ina/PCRS_Comb.seq:*
- 8: /EMC_Celerra_SIDS3/podata/2/ina/PP_COMB.seq:*
- 9: /EMC_Celerra_SIDS3/podata/2/ina/RE_Comb.seq:*
- 10: /EMC_Celerra_SIDS3/podata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB ID | Description |
|------------|--------|-------|--------|--------|---------------------|
| 1 | 2155.2 | 61.2 | 2318 | 3 | US-09-748-796A-6328 |
| 2 | 423.8 | 12.0 | 507 | 3 | US-09-796A-6327 |
| C | 3 | 168.2 | 4.8 | 537 | Sequence 6328, AP |
| C | 4 | 118.2 | 3.4 | 2445 | Sequence 6327, AP |
| C | 5 | 116.4 | 3.3 | 688 | Sequence 6328, AP |
| C | 6 | 104.2 | 3.0 | 575 | Sequence 6327, AP |
| C | 7 | 74.4 | 2.1 | 50000 | Sequence 6328, AP |
| C | 8 | 73.4 | 2.1 | 708 | Sequence 6327, AP |
| C | 9 | 71.8 | 2.0 | 19124 | Sequence 6328, AP |
| C | 10 | 70.4 | 2.0 | 662 | Sequence 6328, AP |
| C | 11 | 70.2 | 2.0 | 1651 | Sequence 6328, AP |
| C | 12 | 70.2 | 2.0 | 1682 | Sequence 6328, AP |
| C | 13 | 68.8 | 2.0 | 1141 | Sequence 6328, AP |
| C | 14 | 68.2 | 1.9 | 50000 | Sequence 6328, AP |
| C | 15 | 68 | 1.9 | 76767 | Sequence 6328, AP |
| C | 16 | 68 | 1.9 | 76767 | Sequence 6328, AP |
| C | 17 | 67.6 | 1.9 | 2917 | Sequence 6328, AP |
| C | 18 | 67.6 | 1.9 | 3322 | Sequence 6328, AP |
| C | 19 | 67.6 | 1.9 | 60376 | Sequence 6328, AP |
| C | 20 | 66.8 | 1.9 | 1410 | Sequence 6328, AP |
| C | 21 | 66.8 | 1.9 | 14066 | Sequence 6328, AP |
| C | 22 | 66.4 | 1.9 | 19438 | Sequence 6328, AP |
| C | 23 | 66 | 1.9 | 26000 | Sequence 6328, AP |
| C | 24 | 65.8 | 1.9 | 205044 | Sequence 6328, AP |
| C | 25 | 65.8 | 1.9 | 205044 | Sequence 6328, AP |
| C | 26 | 65.8 | 1.9 | 205044 | Sequence 6328, AP |
| C | 27 | 65.8 | 1.9 | 223471 | Sequence 6328, AP |
| C | 28 | 65.8 | 1.9 | 223471 | Sequence 6328, AP |
| C | 29 | 65.8 | 1.9 | 223471 | Sequence 6328, AP |

RESULT 1

US-09-248-796A-6328

; Sequence 6328, Application US/09248796A

; Patent No. 6747137

; GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS

; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107116_132

; CURRENT FILING NUMBER: US/09/248-796A

; FILING DATE: 1999-02-12

; PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13

; PRIOR PUBLISHING DATE: 1998-08-13

; NUMBER OF SEQ ID NOS: 2208

; SEQ ID NO 6328

; LENGTH: 2358

; TYPE: DNA

; ORGANISM: Candida albicans

US-09-248-796A-6328

Query Match 61.2%; Score 2155.2%; DB 3; Length 2358;

Best Local Similarity 97.2%; Pred. No. 0;

Mismatches 0; Conservative 0; Indels 44; Gaps 8;

Query Match 61.2%; Score 2155.2%; DB 3; Length 2358;

Best Local Similarity 97.2%; Pred. No. 0;

Mismatches 2293; Conservative 0; Indels 44; Gaps 8;

Query Match 61.2%; Score 2155.2%; DB 3; Length 2358;

Best Local Similarity 97.2%; Pred. No. 0;

Mismatches 2293; Conservative 0; Indels 44; Gaps 8;

Query Match 61.2%; Score 2155.2%; DB 3; Length 2358;

Best Local Similarity 97.2%; Pred. No. 0;

Mismatches 2293; Conservative 0; Indels 44; Gaps 8;

Query Match 61.2%; Score 2155.2%; DB 3; Length 2358;

Best Local Similarity 97.2%; Pred. No. 0;

Mismatches 2293; Conservative 0; Indels 44; Gaps 8;

Query Match 61.2%; Score 2155.2%; DB 3; Length 2358;

Best Local Similarity 97.2%; Pred. No. 0;

Mismatches 2293; Conservative 0; Indels 44; Gaps 8;

Query Match 61.2%; Score 2155.2%; DB 3; Length 2358;

Best Local Similarity 97.2%; Pred. No. 0;

Mismatches 2293; Conservative 0; Indels 44; Gaps 8;

Query Match 61.2%; Score 2155.2%; DB 3; Length 2358;

Best Local Similarity 97.2%; Pred. No. 0;

Mismatches 2293; Conservative 0; Indels 44; Gaps 8;

Query Match 61.2%; Score 2155.2%; DB 3; Length 2358;

Best Local Similarity 97.2%; Pred. No. 0;

Mismatches 2293; Conservative 0; Indels 44; Gaps 8;

Query Match 61.2%; Score 2155.2%; DB 3; Length 2358;

Best Local Similarity 97.2%; Pred. No. 0;

Mismatches 2293; Conservative 0; Indels 44; Gaps 8;

Query Match 61.2%; Score 2155.2%; DB 3; Length 2358;

Best Local Similarity 97.2%; Pred. No. 0;

Mismatches 2293; Conservative 0; Indels 44; Gaps 8;

Query Match 61.2%; Score 2155.2%; DB 3; Length 2358;

Best Local Similarity 97.2%; Pred. No. 0;

Mismatches 2293; Conservative 0; Indels 44; Gaps 8;

Query Match 61.2%; Score 2155.2%; DB 3; Length 2358;

Best Local Similarity 97.2%; Pred. No. 0;

Mismatches 2293; Conservative 0; Indels 44; Gaps 8;

Query Match 61.2%; Score 2155.2%; DB 3; Length 2358;

Best Local Similarity 97.2%; Pred. No. 0;

Mismatches 2293; Conservative 0; Indels 44; Gaps 8;

ALIGNMENTS

| Sequence Data for 1619 AAGAGTGTGATTATTAAAGTTTATTCATCTACTTTCAGTCGATA 1671 | | | | | | | | | |
|--|----|------|---|------|--|--|--|--|--|
| Db | Dy | 753 | ATCATACTCATCTGCAATTCCATTCAACCCATTCAATTCTGAACTTACTAATG | 812 | | | | | |
| Db | Dy | 479 | ATCATACTCATCTGCAATTCCATTCAACCCATTCAATTCTGAACTTACTAATG | 538 | | | | | |
| Db | Dy | 813 | CTDATAATAATAGAAGAAAATTCTCAATTGAAATTAGCTCATTAAAGAAA | 872 | | | | | |
| Db | Dy | 539 | CTATAATACAGAGAAATTCTCAATTGAAATTAGCTCATTAAAGAAA | 598 | | | | | |
| Db | Dy | 873 | CATTGTCCTTAATTGAAATTAGCTCATTAAAGAAA | 932 | | | | | |
| Db | Dy | 599 | CATTGTCCTTAATTGAAATTAGCTCATTAAAGAAA | 658 | | | | | |
| Db | Dy | 933 | TGAAACTATTGAAATTGAGAAATTCTGAAATTCTGAAATTCTGAAATTCTGAA | 989 | | | | | |
| Db | Dy | 659 | TGAAACTATTGAAATTCTGAAATTCTGAAATTCTGAAATTCTGAAATTCTGAA | 718 | | | | | |
| Db | Dy | 990 | AACCTGGTATGTCGTCATCCACCCGCACTGCGCTACTGCGCTACTGCG | 1049 | | | | | |
| Db | Dy | 790 | AACCTGGTATGTCGTCATCCACCCGCACTGCGCTACTGCGCTACTGCG | 1160 | | | | | |
| Db | Dy | 719 | AACCTGGTATGTCGTCATCCACCCGCACTGCGCTACTGCGCTACTGCG | 778 | | | | | |
| Db | Dy | 1050 | CCACTGCTGTC-----TGAGACCATCGTCATCTATTGATGAG | 1100 | | | | | |
| Db | Dy | 779 | CCACTGCTGTCATCTGTCGTCATCTGTCGTCATCTGTCGTCATCTGTC | 838 | | | | | |
| Db | Dy | 1101 | CTCATCTCATGTCATCAAGATCATCTCAGAACTTCAGAACTTCAGAA | 839 | | | | | |
| Db | Dy | 839 | CTGATCTCATGTCATCAAGATCATCTCAGAACTTCAGAACTTCAGAA | 958 | | | | | |
| Db | Dy | 1161 | TTCGTGATCATATTAGTGTATTGGAATTGATGAAATTGATGAAATTGATG | 1220 | | | | | |
| Db | Dy | 899 | TTCGTGATCATATTAGTGTATTGATGAAATTGATGAAATTGATGAAATTGATG | 958 | | | | | |
| Db | Dy | 1221 | AGAGAGAGAGAGCTTATTCACAAACAACTAACATGCTGATCTGCTATTG | 1280 | | | | | |
| Db | Dy | 959 | AGAGAGAGAGAGCTTATTCACAAACAACTAACATGCTGATCTGCTATTG | 1018 | | | | | |
| Db | Dy | 1281 | ATGAAATTCAAAATCTTCAAAAAATAACTCTAGTGCGCTCATCGCTCAT | 1340 | | | | | |
| Db | Dy | 1019 | ATGAAATTCAAAATCTTCAAAAAATAACTCTAGTGCGCTCATCGCTCAT | 1078 | | | | | |
| Db | Dy | 1341 | ATCACTACGTTATTAATAAAATTTGTTGTTGATGTTGTTGTTGTTGATG | 1397 | | | | | |
| Db | Dy | 1079 | ATCACTACGTTATTAATAAAATTTGTTGTTGATGTTGTTGTTGTTGATG | 1138 | | | | | |
| Db | Dy | 1398 | TGCCAGCTTAAATAATCTCCAAAATAATTCTGAACTTACGAACTTACGAA | 1457 | | | | | |
| Db | Dy | 1139 | TGCCAGCTTAAATAATCTCCAAAATAACTCTGAACTTACGAACTTACGAA | 1198 | | | | | |
| Db | Dy | 1458 | TTACTCATGAAATTCTGAAATTCTGAAATTCTGAAATTCTGAAATTCTGAA | 1517 | | | | | |
| Db | Dy | 1199 | TTACTCATGAAATTCTGAAATTCTGAAATTCTGAAATTCTGAAATTCTGAA | 1258 | | | | | |
| Db | Dy | 1518 | CACCTTCATAATTATGATCTATTCTGAAATTCTGAAATTCTGAAATTCTGAA | 1571 | | | | | |
| Db | Dy | 1259 | CACCTTCATAATTATGATCTATTCTGAAATTCTGAAATTCTGAAATTCTGAA | 1118 | | | | | |
| Db | Dy | 1572 | CGGATCTGCACTGGCAGGGAAATTGGCAGGAAATTCTGAAATTCTGAAATTG | 1631 | | | | | |
| Db | Dy | 1319 | CGGATCTGCACTGGCAGGGAAATTGGCAGGAAATTCTGAAATTCTGAAATTG | 1378 | | | | | |
| Db | Dy | 1632 | GAGGTCTTACCTGATGATTCTCATTTCTGAAATTCTGAAATTCTGAAATTG | 1691 | | | | | |
| Db | Dy | 1379 | GAGGTCTTACCTGATGATTCTCATTTCTGAAATTCTGAAATTCTGAAATTG | 1438 | | | | | |
| Db | Dy | 1692 | CCCCGGATATTCCATGTCATGTCATGTCATGTCATGTCATGTCATGTC | 1551 | | | | | |
| Db | Dy | 1439 | CCCCGGATATTCCATGTCATGTCATGTCATGTCATGTCATGTCATGTC | 1498 | | | | | |
| Db | Dy | 1752 | GTCGAGAACTGGCTGCTGTTGATGTTGATGTTGATGTTGATGTTGATG | 1871 | | | | | |
| Db | Dy | 1559 | GTCGAGAACTGGCTGCTGTTGATGTTGATGTTGATGTTGATGTTGATG | 1618 | | | | | |
| Db | Dy | 1499 | GTCGAGAACTGGCTGCTGTTGATGTTGATGTTGATGTTGATGTTGATG | 1558 | | | | | |
| Db | Dy | 1812 | CCAAAGATTCTGATTCTGATTCTGAAATTCTGAAATTCTGAAATTCTGAA | 1811 | | | | | |
| Db | Dy | 1559 | CCAAAGATTCTGATTCTGAAATTCTGAAATTCTGAAATTCTGAAATTCTGAA | 1618 | | | | | |
| Db | Dy | 1872 | AAAAAGTGTGATTATTAAAGTTTATTCATCTGTTGTTGTTGTTGTTGTTG | 1931 | | | | | |

GenCore version 5.1.9
(c) 1993 - 2006 Bioceleration Ltd.

Om protein - protein search, using sw model

Run on: June 13, 2006, 14:21:02 : Search time 17 Seconds

(without alignments)
685.7/08 Million cell updates/sec

Title: US-10-018-105A-4

Perfect score: 4820

Sequence: 1. MSDSBYTQNNTNQPIRS. LGIKKHHGGNNKTSISPPNKYE 922

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 64916 seqs, 12643261 residues

Total number of hits satisfying chosen parameters: 64916

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_New.*

1: /EMC_CeIerra_SIDS3/picodata/1/pubpbaa/US02_NEW_PUB_pep: *
2: /EMC_CeIerra_SIDS3/picodata/1/pubpbaa/US05_NEW_PUB_pep: *
3: /EMC_CeIerra_SIDS3/picodata/1/pubpbaa/US07_NEW_PUB_pep: *
4: /EMC_CeIerra_SIDS3/picodata/1/pubpbaa/US08_NEW_PUB_pep: *
5: /EMC_CeIerra_SIDS3/picodata/1/pubpbaa/PCT_NEW_PUB_pep: *
6: /EMC_CeIerra_SIDS3/picodata/1/pubpbaa/US10_NEW_PUB_pep: *
7: /EMC_CeIerra_SIDS3/picodata/1/pubpbaa/US11_NEW_PUB_pep: *
8: /EMC_CeIerra_SIDS3/picodata/1/pubpbaa/US60_NEW_PUB_pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query | Match | Length | DB ID | Description |
|------------|-------|-------|-------|--------|---------------------|-------------------|
| 1 | 166.5 | 3.5 | 2271 | 6 | US-10-018-105A-4 | Sequence 3948, AP |
| 2 | 157 | 3.3 | 750 | 7 | US-11-013-228A-32 | Sequence 32, Appl |
| 3 | 137 | 2.8 | 454 | 7 | US-11-013-228A-35 | Sequence 35, Appl |
| 4 | 137 | 2.8 | 1392 | 6 | US-10-018-349-1450 | Sequence 1410, AP |
| 5 | 116 | 2.8 | 315 | 6 | US-10-018-571A-484 | Sequence 334, AP |
| 6 | 136 | 2.8 | 2478 | 6 | US-10-018-571A-278 | Sequence 2278, AP |
| 7 | 135.5 | 2.8 | 596 | 6 | US-10-018-749-310 | Sequence 100, AP |
| 8 | 135.5 | 2.8 | 933 | 7 | US-11-011-316-100 | Sequence 178, AP |
| 9 | 133 | 2.8 | 933 | 7 | US-11-012-154-178 | Sequence 898, AP |
| 10 | 130 | 2.7 | 1274 | 6 | US-10-071-571A-888 | Sequence 3, Appl |
| 11 | 129.5 | 2.7 | 624 | 6 | US-10-028-104-3 | Sequence 10, Appl |
| 12 | 128.5 | 2.7 | 486 | 6 | US-10-018-571A-10 | Sequence 4150, AP |
| 13 | 128.5 | 2.7 | 788 | 7 | US-11-013-697-4150 | Sequence 33, Appl |
| 14 | 126 | 2.6 | 710 | 7 | US-11-013-228A-34 | Sequence 34, Appl |
| 15 | 124.5 | 2.6 | 477 | 7 | US-11-013-228A-34 | Sequence 4496, AP |
| 16 | 124.5 | 2.6 | 9535 | 6 | US-10-071-571A-4496 | Sequence 2758, AP |
| 17 | 123 | 2.6 | 688 | 7 | US-11-013-697-2798 | Sequence 3781, AP |
| 18 | 121 | 2.5 | 940 | 7 | US-11-013-697-3791 | Sequence 3781, AP |
| 19 | 119.5 | 2.5 | 774 | 6 | US-10-018-571A-484 | Sequence 4844, AP |
| 20 | 119.5 | 2.5 | 791 | 7 | US-11-011-154-24 | Sequence 24, Appl |
| 21 | 119.5 | 2.5 | 1106 | 7 | US-11-013-697-4301 | Sequence 4301, AP |
| 22 | 119 | 2.5 | 502 | 6 | US-10-011-937-3008 | Sequence 3008, AP |
| 23 | 118 | 2.4 | 387 | 6 | US-10-053-349-8568 | Sequence 8568, AP |
| 24 | 118 | 2.4 | 439 | 6 | US-10-053-349-8567 | Sequence 8567, AP |
| 25 | 117 | 2.4 | 1003 | 6 | US-10-053-349-1411 | Sequence 1411, AP |
| 26 | 116.5 | 2.4 | 2937 | 6 | US-10-018-962-18 | Sequence 18, Appl |
| 27 | 116.5 | 2.4 | 2969 | 6 | US-10-018-962-19 | Sequence 19, Appl |
| 28 | 116 | 2.4 | 1075 | 7 | US-11-013-228A-23 | Sequence 23, Appl |
| 29 | 115.5 | 2.4 | 1261 | 6 | US-10-071-571A-1304 | Sequence 1304, AP |
| 30 | 113 | 2.3 | 611 | 6 | US-10-053-349-3922 | Sequence 3922, AP |
| 31 | 113 | 2.3 | 6214 | 6 | US-10-053-349-3921 | Sequence 3921, AP |

ALIGNMENTS

| RESULT 1 | | | | | | | | | |
|------------------|-------------|---------------------------|----------------|----------------------------------|----------------------|-----------|--|----------------------------------|--|
| US-10-018-105A-4 | Score: 3948 | Application: US/10471571A | Sequence: 3948 | Publication No.: US20060115490A1 | General Information: | Priority: | Priority Application Number: US/10/471571A | Priority Filing Date: 2003-09-12 | Prior Art Application Number: GB-0107661.1 |
| | | | | | | | | | SEQ ID NO: 562 |
| | | | | | | | | | Software: SeqWin99, version 1.03 |
| | | | | | | | | | Length: 2271 |
| | | | | | | | | | Type: PRT |
| | | | | | | | | | Organism: Staphylococcus aureus |
| | | | | | | | | | Feature: Feature |
| | | | | | | | | | NAME KEY: MISC_FEATURE |
| | | | | | | | | | LOCATION: (1)..(2271) |
| | | | | | | | | | OTHER INFORMATION: hypothetical protein |
| | | | | | | | | | US-10-018-105A-4 |

| Query | Match | Score | DB | Length | Best Local Similarity | Matches | Conservative | Mismatches | Indels | Gaps |
|-------|-------|--|------|--------|-----------------------|---------|--------------|------------|--------|------|
| Qy | 1 | MSDSEBYTQNNTNQPIRSDEVDLDRHQINNDTCAISDSEDELKLSELESEVVKSKQ | 60 | 20 | 3.5% | 3 | 18.3% | 0 | 0.052 | 0 |
| Qy | 1 | MSDSEBYTQNNTNQPIRSDEVDLDRHQINNDTCAISDSEDELKLSELESEVVKSKQ | 60 | 20 | 3.5% | 3 | 18.3% | 0 | 0.052 | 0 |
| Db | 1701 | ISDQMSSESVDNSSEVSNSSESDS-GSLSGSPTSVSLRKSESVS | 1757 | 20 | 3.5% | 3 | 18.3% | 0 | 0.052 | 0 |
| Qy | 61 | QQHQHETTSNDNAPLTRKS---GSISKKSNLTDKDRTNPMSLGGDDTINSCHKRN | 117 | 20 | 3.5% | 3 | 18.3% | 0 | 0.052 | 0 |
| Db | 1758 | SLSCSQNSDSVTSFSDSSLSVTSLSRSSEVSSTSDSLSKTSGSTSSTSLSLST | 1817 | 20 | 3.5% | 3 | 18.3% | 0 | 0.052 | 0 |
| Qy | 118 | NMS---SLRDKYTKLN---TDDNNTNNTHLAIPPIPPIPIT-----NANSR | 163 | 20 | 3.5% | 3 | 18.3% | 0 | 0.052 | 0 |
| Db | 1818 | SLSGSESVSSTSDSLSKTSGSTSSTSLSLSDSMDSTSDSLSLSSQ | 1877 | 20 | 3.5% | 3 | 18.3% | 0 | 0.052 | 0 |
| Qy | 164 | RKSGLENPLPLKKIGRNNSPENLDV--SPMTKAKTSDSNTSTAAHRLG- | 220 | 20 | 3.5% | 3 | 18.3% | 0 | 0.052 | 0 |
| Db | 1878 | SMGMSASTSSQTSVTSLSDSDSQTSSNQPSMISASBDSMSTS | 1937 | 20 | 3.5% | 3 | 18.3% | 0 | 0.052 | 0 |
| Qy | 221 | --IGATLGYGTGTTT-----ATATAAAGR----- | 244 | 20 | 3.5% | 3 | 18.3% | 0 | 0.052 | 0 |
| Db | 1938 | SMGGSVSVSTSLSDSDSISGVSSDSSSTSLSLSDSMSQSQTSSCAGLCLSTSLST | 1997 | 20 | 3.5% | 3 | 18.3% | 0 | 0.052 | 0 |
| Qy | 245 | --RPSSSESDADS---HASASQSTEEDVFPAGDHIRVNG | 283 | 20 | 3.5% | 3 | 18.3% | 0 | 0.052 | 0 |
| Db | 1998 | SMSMSASTSSQTSVTSLSDSDSQTSSNQPSMISASBDSMSTS | 2057 | 20 | 3.5% | 3 | 18.3% | 0 | 0.052 | 0 |
| Qy | 284 | IDPDEIDEFTREREPAYLQKOMIAKILRIDEPOLSKNNTISASHPYTHHISNNNNK | 343 | 20 | 3.5% | 3 | 18.3% | 0 | 0.052 | 0 |
| Db | 2058 | SDPS----- | 2140 | 20 | 3.5% | 3 | 18.3% | 0 | 0.052 | 0 |
| Qy | 344 | NGDGCGGSMAALKYTPKNUKLTURPFETHNSSSBELYKTMQPKYDQLS | 403 | 20 | 3.5% | 3 | 18.3% | 0 | 0.052 | 0 |
| Db | 2084 | STSGSSSTSVSDTSMSDKSDSTSISDVSSS | 2140 | 20 | 3.5% | 3 | 18.3% | 0 | 0.052 | 0 |

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2006, 14:08:32 : Search time 52 Seconds

(without alignments)
1551.985 Million cell updates/sec

Title: US-10-018-105A-4

Perfect score: 4820

Sequence: 1 MSDESYTQNNTNQPIPPS.....LGKXHGNNKTSISFPNPKY 922

Scoring table: BLOSUM62

Gapop 10.0 . Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 0%

Listing first 45 summaries

Database : Issued Patents_AA:*

1: /EMC_Celerra_SIDS3_ptodata/2/iaas/5_COMB.pep:*

2: /EMC_Celerra_SIDS3_ptodata/2/iaas/6_COMB.pep:*

3: /EMC_Celerra_SIDS3_ptodata/2/iaas/7_COMB.pep:*

4: /EMC_Celerra_SIDS3_ptodata/2/iaas/8_COMB.pep:*

5: /EMC_Celerra_SIDS3_ptodata/2/iaas/9_COMB.pep:*

6: /EMC_Celerra_SIDS3_ptodata/2/iaas/10_COMB.pep:*

7: /EMC_Celerra_SIDS3_ptodata/2/iaas/RE_COMB.pep:*

7: /EMC_Celerra_SIDS3_ptodata/2/iaas/backfilesl1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result | No. | Score | Query | Match | Length | DB | ID | Description |
|--------|-------|-------|-------|-------|----------------------|--------------------|----|-------------|
| 1 | 3785 | 78.5 | 785 | 2 | US-09-248-796A-0431 | Sequence 20431, A | | |
| 2 | 1680 | 34.9 | 659 | 2 | US-09-533-092-171 | Sequence 717, App | | |
| 3 | 834 | 17.3 | 814 | 2 | US-09-248-796A-20405 | Sequence 20405, A | | |
| 4 | 716 | 14.9 | 168 | 2 | US-09-248-796A-20430 | Sequence 20430, A | | |
| 5 | 251 | 5.2 | 326 | 2 | US-09-902-540-11640 | Sequence 11640, A | | |
| 6 | 172.5 | 3.6 | 1091 | 2 | US-09-538-092-66 | Sequence 469, App | | |
| 7 | 166.5 | 3.5 | 2283 | 2 | US-10-177-502-4 | Sequence 4, App1 | | |
| 8 | 166 | 3.4 | 719 | 2 | US-09-248-796A-17559 | Sequence 17559, A | | |
| 9 | 161 | 3.3 | 366 | 2 | US-09-712-363-193 | Sequence 193, App | | |
| 10 | 161 | 3.3 | 471 | 2 | US-08-311-731A-168 | Sequence 168, App | | |
| 11 | 159.5 | 3.3 | 1177 | 2 | US-09-134-001C-5106 | Sequence 5106, App | | |
| 12 | 159.5 | 3.3 | 1253 | 2 | US-09-958-617A-18 | Sequence 18, App1 | | |
| 13 | 159.5 | 3.3 | 1584 | 2 | US-09-457-040B-27 | Sequence 27, App1 | | |
| 14 | 159.5 | 3.3 | 2150 | 3 | US-10-115-322-17 | Sequence 17, App1 | | |
| 15 | 158.5 | 3.3 | 785 | 2 | US-09-481-538B-348 | Sequence 348, App | | |
| 16 | 158 | 3.3 | 651 | 1 | US-08-431-050-24 | Sequence 24, App1 | | |
| 17 | 158 | 3.3 | 651 | 1 | US-08-938-534-24 | Sequence 24, App1 | | |
| 18 | 158 | 3.3 | 651 | 2 | US-09-345-294-24 | Sequence 24, App1 | | |
| 19 | 157 | 3.3 | 858 | 2 | US-09-248-796A-19055 | Sequence 19055, A | | |
| 20 | 154 | 3.2 | 1259 | 2 | US-09-949-016-10366 | Sequence 10366, A | | |
| 21 | 151.5 | 3.2 | 982 | 2 | US-09-248-796A-0628 | Sequence 20628, A | | |
| 22 | 152 | 3.2 | 2870 | 2 | US-09-479-467A-15 | Sequence 15, App1 | | |
| 23 | 152 | 3.2 | 3178 | 2 | US-09-655-160-15 | Sequence 15, App1 | | |
| 24 | 152 | 3.2 | 3178 | 2 | US-09-479-467A-4 | Sequence 4, App1 | | |
| 25 | 152 | 3.2 | 3178 | 2 | US-09-655-160-4 | Sequence 4, App1 | | |
| 26 | 151 | 3.1 | 315 | 2 | US-09-710-279-2100 | Sequence 2100, AD | | |
| 27 | 151 | 3.1 | 324 | 2 | US-09-134-001C-4080 | Sequence 4080, AD | | |
| 28 | 151 | 3.1 | 10182 | 2 | US-09-134-001C-1319 | Sequence 315, AD | | |
| 29 | 150 | 3.1 | 676 | 2 | US-09-134-001C-4118 | Sequence 4318, AD | | |
| 30 | 147 | 3.0 | 404 | 2 | US-09-710-279-398 | Sequence 398, AD | | |
| 31 | 147 | 3.0 | 5024 | 2 | US-09-710-279-2964 | Sequence 2962, AD | | |
| 32 | 146.5 | 3.0 | 1444 | 2 | US-09-949-016-9652 | Sequence 9652, AD | | |

RESULT 1

US-09-248-796A-20431

1 Sequence 20431, Application US/09248796A

2 Patent No. 6747137

3 GENERAL INFORMATION:

4 APPLICANT: Keith Weinstock et al

5 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA

6 ALBICANS

7 FILE REFERENCE: 107196_132

8 CURRENT APPLICATION NUMBER: US/09/248-796A

9 CURRENT PUBLISHING DATE: 1998-02-12

10 PRIOR APPLICATION NUMBER: US 60/074_725

11 PRIOR FILING DATE: 1998-02-13

12 PRIOR APPLICATION NUMBER: US 60/096_409

13 NUMBER OF SEQ ID NOS: 28208

14 SEQ ID NO: 28208

15 LENGTH: 785

16 TYPE: PRT

17 ORGANISM: Candida albicans

18 US-09-248-796A-20431

19 Query Match 78.5%; Score 3785; DB 2; Length 785;

20 Best Local Similarity 96.2%; Pred. No. 3e-996; 2; Mismatches 13; Indels 14; Caps 7;

21 Qy 1 MDSSESSIONSTRQIPTSDVCDHNRQITDCAISDSDLELSEVEVKSE-59

22 Db 18 MSDSBTSYNTNSTNPILRSDEVDLDRNQNTDCAISDSDLELKPELESAEKQ 77

23 Qy 60 ---QQQHQETTSNDNAKLTQKSGSSTKKSNLTDKRTINPNLSGGDDTTNSGKRN 116

24 Db 78 QQQQHQETTSNDNAKLTQKSGSSTKKSNLTDKRTINPNLSGGDDTTNSGKRN 137

25 Qy 117 YMSSLRDFLKLDTNDSTNTHAIPPIPTPITNTANKSRKQLSLENLPPLIK 176

26 Db 138 YMSSLRDFLKLDTNDSTNTHAIPPIPTPITNTANKSRKQLSLENLPPLIK 197

27 Qy 177 KTTGRNNNNPENDLYSPMTKMTNDSEDTN-TSTTANHMKLGATPTLGNGCTTA 235

28 Db 198 KTTGRNNNNPENDLYSPTRATNUEDNTTSTYPAHFLGATPTLGNGCTTA 257

29 Qy 234 ATATAAA ---GRRPSRSSIDEADSHASRSQTEBEDCPAYGDIRYNGIDFEDIDP 292

30 Db 258 ATATATAGGGRRPSRSSIDEADSHASRSQTEBEDACCPAYGDIRYNGIDFEDIDP 317

31 Qy 293 IRRERAYLOKONIAKNTLRIEPONLSKNTTCAASRPYHMKNNKKANGEDS-GG 351

32 Db 318 IRRERAYLQRMLAKNTLRIEPONLSKNTTGAHRPHYHNNNNKINGGGCG 377

33 Qy 332 SSMALKYTPRNKNIKLTSLPFTHEINSSSEYLKTKQOPPYDQYDQLSLTSST 411

34 Db 378 SSMALKYTPRNKNIKLTSLPFTHEINSSSEYLKTKQOPPYDQYDQLSLTSST 437

35 Qy 412 ---SGSGSQWFGARISDGINSQGSPQSTPDPFLVSQSPQSTPDPFLVSQ 469

36 Db 438 SGSGSQWFGARISDGINSQGSPQSTPDPFLVSQSPQSTPDPFLVSQ 497

37 Qy 470 RNGEETWMIDCTPDTSEMCLAKAFCHPLAEFRKVEVLPKSYFYCPHTF 529

GenCore version 5.1.9
 Copyright (c) 1993 - 2006 Biacceleration Ltd.

CM protein - protein search, using sw model

Run on: June 13, 2006, 14:20:38 : Search time 182 Seconds
 (Without alignments)
 2346,616 Millsec updates/sec

Title: US-10-018-105A-4

Perfect score: 4820

Sequence: 1 MSDSEYYQNSTTNPPIRS.....LGKKGGENKTSISPPNKE 922

Scoring table: BLOSUM62

Gapext 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 10%

Listing first 45 summaries

Database : Published_Applications_AA_Main;*

1: /EMC_Celerra_SIDS1/pcodata/2/pubpaa/US07_PUBCOMB.pep;*

2: /EMC_Celerra_SIDS3/pcodata/2/pubpaa/US08_PUBCOMB.pep;*

3: /EMC_Celerra_SIDS3/pcodata/2/pubpaa/US10_PUBCOMB.pep;*

4: /EMC_Celerra_SIDS3/pcodata/2/pubpaa/US10B_PUBCOMB.pep;*

5: /EMC_Celerra_SIDS3/pcodata/2/pubpaa/US11_PUBCOMB.pep;*

6: /EMC_Celerra_SIDS1/pcodata/2/pubpaa/US11_PUBCOMB.pep;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query | Match | Length | DB ID | Description |
|------------|-------|-------|-------|--------|---------------------|---------------------|
| 1 | 4814 | 99.9 | 922 | 4 | US-10-012-585-7675 | Sequence 7675, AP |
| 2 | 1679 | 34.8 | 859 | 2 | US-0-945-749-1 | Sequence 2, Appli |
| 3 | 1666 | 34.6 | 858 | 2 | US-0-945-749-1 | Sequence 2, Appli |
| 4 | 1059 | 22.8 | 663 | 4 | US-10-128-714-3563 | Sequence 8563, AP |
| 5 | 1059 | 22.8 | 663 | 4 | US-10-128-714-3563 | Sequence 8563, AP |
| 6 | 856 | 17.8 | 969 | 2 | US-0-945-749-3 | Sequence 3, Appli |
| 7 | 3455 | 7.2 | 154 | 5 | US-0-945-749-3 | Sequence 2166, AP |
| 8 | 298 | 6.2 | 94 | 4 | US-10-45-11-321932 | Sequence 321932, AP |
| 9 | 234 | 6.2 | 387 | 2 | US-0-945-749-8 | Sequence 8, Appli |
| 10 | 234 | 4.9 | 1198 | 6 | US-11-097-143-7905 | Sequence 7905, AP |
| 11 | 213 | 4.4 | 380 | 2 | US-0-945-749-7 | Sequence 8, Appli |
| 12 | 213 | 4.4 | 386 | 4 | US-10-156-761-9574 | Sequence 9574, AP |
| 13 | 201 | 4.2 | 371 | 4 | US-10-156-761-1064 | Sequence 1064, A |
| 14 | 182 | 3.9 | 373 | 3 | US-0-945-749-5 | Sequence 1568, AP |
| 15 | 179 | 3.5 | 373 | 6 | US-11-097-143-7905 | Sequence 48, AP |
| 16 | 179 | 3.7 | 1172 | 4 | US-10-451-677-574 | Sequence 574, AP |
| 17 | 179 | 3.7 | 2344 | 5 | US-10-122-223-13978 | Sequence 13978, A |
| 18 | 179 | 3.7 | 354 | 6 | US-10-156-761-13791 | Sequence 13791, A |
| 19 | 179 | 3.7 | 1199 | 6 | US-11-097-143-1614 | Sequence 1614, AP |
| 20 | 175 | 3.6 | 1246 | 6 | US-11-097-143-1614 | Sequence 1143, A |
| 21 | 172 | 3.6 | 3328 | 5 | US-10-732-923-8311 | Sequence 8311, AP |
| 22 | 172 | 3.6 | 2165 | 5 | US-10-732-93-31347 | Sequence 13347, A |
| 23 | 170 | 3.5 | 343 | 4 | US-10-156-761-14171 | Sequence 14171, A |
| 24 | 169 | 3.5 | 1468 | 6 | US-11-097-143-22779 | Sequence 22779, A |
| 25 | 168 | 3.5 | 816 | 5 | US-10-457-555-32 | Sequence 32, Appli |
| 26 | 166 | 3.5 | 1046 | 6 | US-11-165-819-6 | Sequence 6, Appli |
| 27 | 166 | 3.5 | 1046 | 6 | US-11-165-819-6 | Sequence 26, Appli |
| 28 | 166 | 3.5 | 2261 | 5 | US-10-410-048-60 | Sequence 60, Appli |
| 29 | 166 | 3.5 | 2283 | 4 | US-10-127-502-4 | Sequence 4, Appli |
| 30 | 166 | 3.5 | 2283 | 6 | US-11-025-509-4 | Sequence 4, Appli |
| 31 | 166 | 3.4 | 688 | 4 | US-10-352-585-7876 | Sequence 7876, AP |
| 32 | 165 | 3.4 | 2112 | 6 | US-11-097-143-8001 | Sequence 8001, AP |
| 33 | 165 | 3.4 | 4498 | 4 | US-10-712-124-68 | Sequence 68, Appli |

ALIGNMENTS

| RESULT 1 | | | | | | |
|---|--|--|--|--|--|--|
| US-10-032-585-7675 | | | | | | |
| : Sequence 7675, Application US/10032585 | | | | | | |
| : Publication No. US2003180953A1 | | | | | | |
| : GENERAL INFORMATION: | | | | | | |
| : APPLICANT: Perry, Roger D. | | | | | | |
| : APPLICANT: Bo, Jiang | | | | | | |
| : APPLICANT: Charles, Boone | | | | | | |
| : APPLICANT: Howard, Bussey | | | | | | |
| : TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery | | | | | | |
| : FILE REFERENCE: 10182-005-939 | | | | | | |
| : CURRENT APPLICATION NUMBER: US10/032,585 | | | | | | |
| : CURRENT FILING DATE: 2001-12-20 | | | | | | |
| : NUMBER OF SEQ ID NOS: 000 | | | | | | |
| : SOFTWARE: Patentin version 3.1 | | | | | | |
| : SEQ ID NO: 7675 | | | | | | |
| : LENGTH: 922 | | | | | | |
| : TYPE: PRT | | | | | | |
| : ORGANISM: Candida albicans | | | | | | |
| US-10-032-585-7675 | | | | | | |
| Query Match 99.9%; Score 4814; DB 4; Length 922; | | | | | | |
| Best Local Similarity 99.9%; Pred. No: 0; | | | | | | |
| Matches 921; Conservative 0; Mismatches 1; Indels 0; Gaps 0; | | | | | | |
| Qy 1 MSDSEYYQNSTTNPPIRS.....LGKKGGENKTSISPPNKE 922 | | | | | | |
| Db 1 MSDSEYYQNSTTNPPIRS.....LGKKGGENKTSISPPNKE 60 | | | | | | |
| Qy 1 MSDSEYYQNSTTNPPIRS.....LGKKGGENKTSISPPNKE 60 | | | | | | |
| Db 1 MSDSEYYQNSTTNPPIRS.....LGKKGGENKTSISPPNKE 60 | | | | | | |
| Qy 61 QHQHQTSDARPLTKSSEIKKSINQDQRITNPMSLGGDDTINSKGKNTYNS 120 | | | | | | |
| Db 61 QHQHQTSDARPLTKSSEIKKSINQDQRITNPMSLGGDDTINSKGKNTYNS 120 | | | | | | |
| Qy 121 SLRDKVLUKNTDDNSTGNGTHIAPIPPIPPIPITMANKSRKSQLENPLPLKKTI 180 | | | | | | |
| Db 121 SLRDKVLUKNTDDNSTGNGTHIAPIPPIPPIPITMANKSRKSQLENPLPLKKTI 180 | | | | | | |
| Qy 181 GRNNNNFENDLYSPMTKNTDSEDNTNTSTANPKMKIGATLGVGTGTATA 240 | | | | | | |
| Db 181 GRNNNNFENDLYSPMTKNTDSEDNTNTSTANPKMKIGATLGVGTGTATA 240 | | | | | | |
| Qy 241 AAGRPRSRSDSIDADSHASRSSOETEDVCPMGDHIVNGIDDEDETREREA 300 | | | | | | |
| Db 241 AAGRPRSRSDSIDADSHASRSSQETEDCFAVEDHIVNGIDDEDETREREA 300 | | | | | | |
| Qy 301 YLQOMTAKIILRIDDEFQUNSKNTSGASRHYYHHHSNNNNKRNQGDGGSSMAALYT 360 | | | | | | |
| Db 301 YLQOMTAKIILRIDDEFQUNSKNTSGASRHYYHHHSNNNNKRNQGDGGSSMAALYT 360 | | | | | | |
| Qy 361 PKNLKTKLSEPFETHEENSSSPPEVETAKTQKRPKYYDOLSLMSNSNSGSGQVK 420 | | | | | | |
| Db 361 PKNLKTKLSEPFETHEENSSSPPEVETAKTQKPPKTYDQLSTSSTSFGSGQVK 420 | | | | | | |
| Qy 421 FGGRISPQDGINGSLPLDRFLPSHSESEBTHADPDIPLSPQSTRLFRNCETWMLDC 480 | | | | | | |
| Db 421 FGGRISPQDGINGSLPLPRFLPHSESEBTHADPDIPLSPQSTRLFRNCETWMLDC 480 | | | | | | |
| Qy 541 INVIVVFHDGILTPHSPSPISHPANTRRVRQLRDYDVSADMLYALIDEITDGFPAVI 600 | | | | | | |

| | | | | | | | |
|----------|--------------------------------|--|-----|----|-----------------|--|-----|
| Db | 541 | INYYIVFHGILTHPSLSPANVRRVRQLRDTDVSDMCLCYALIDEITDPAVI | 600 | Db | 269 | ETEEDVCFPMPOLHTRVNGIDPDELEETAYQFAMAB--KSQFLASLQVNPNEQKYSVQ | 325 |
| Oy | 601 | HGIEYEAIDAAIDAVTARDTDSMLORIGESBRKTMLRLLSGLADYTKNFAKRCOE | 660 | Oy | 323 | --NNTSGASHRYHHHSNNNNKANNNGDCGGSMAALYTPTNLKTKLSPFETH | 376 |
| Db | 601 | HGIEYEAIDAAIDAVTARDTDSMLORIGESBRKTMLRLLSGLADYTKNFAKRCOE | 660 | Db | 326 | DIGFTSTSTSGSS- | |
| Oy | 661 | AMSSGGYQRQYLQOCQQA PPPERPIPLTSPINSTLNLSLCTGGCIVGQNGCP | 720 | Oy | 377 | ENSSSBETELTKOOPPKYDQSLSLTSSTSISGSQGVQKFG-AIESDGINGSSL | 435 |
| Db | 661 | AMSSGGYQRQYLQOCQQA PPPERPIPLTSPINSTLNLSLCTGGCIVGQNGCP | 720 | Db | 360 | ET----BIIHKKDEHEKPSKSHPSISGKRVGEENENIPSDPACSYQCDPQI | 414 |
| Oy | 721 | NPGANTNTNTGSPSPQQQHGHTKNSPPIPDPARADIALYLGDIODHITMFQ | 780 | Oy | 436 | PDRFLPHSSSEETTAHDIPSUSPCQSVRDLPFGEETTWLDCTCPIDSEMOKMALARAP | 495 |
| Db | 721 | NPGANTNTNTGSPSPQQQHGHTKNSPPIPDPARADIALYLGDIODHITMFQ | 780 | Db | 415 | PNRFSFCSCSDEVHSDPSLVSQSPYELPRQCEPTPWLDSCPPTDEMOKMALARAP | 474 |
| Oy | 781 | NLLATEKPSHSMYLAQLOQESPNSNNKIMPEKTLISTMLPMLAULTCGANCW | 840 | Oy | 496 | GHPULTADEMRQPTREKVELPSTYVCPHTPAKEDSYLDPENNTVOPADICUTP | 555 |
| Db | 781 | NLLATEKPSHSMYLAQLOQESPNSNNKIMPEKTLISTMLPMLAULTCGANCW | 840 | Db | 475 | GHPMLTBDRMOTREKVELFSYFVCPHTENDKESEDPLEPINVYTVCRSGVIAUTP | 534 |
| Oy | 841 | VPCEGCTNLMGPGLGVLYLIFTIGSTPAQWKLKANNISBGQNNGRPIPNHSSRSI | 900 | Oy | 556 | HFSPISHPAVNRVRVROLRDYVSDADMICYALIDEITCGPAVPHGELEYADATEDAVP | 615 |
| Db | 841 | VPCEGCTNLMGPGLGVLYLIFTIGSTPAQWKLKANNISBGQNNGRPIPNHSSRSI | 900 | Db | 535 | HFGTISICANBVRVQLDQYVNVNSDNWLCYALIDDITDSFAPVQSYEYADATEDAVP | 594 |
| Oy | 901 | RSLGLKKGKNSLISPNKYE | 922 | Oy | 616 | TARDDFSSMLQRIGESERRAKVTLMLSSKAVIDPARKRQEANSSGYQRQYLNQ | 675 |
| Db | 901 | RSLGLKKGKNSLISPNKYE | 922 | Db | 595 | MARMDDAAUQORGESERKTTMLLSKAVIDPARKRQEANSSGYQRQYLNQ | 675 |
| Oy | 644 | QQQQAPPAPPNPPINTSPINSTLNLSLGTSGGGVG/GV/GTGNPQGNTNTNTTG | 735 | Oy | 643 | -----SHVKNNISSTTPNRY | |
| Db | 644 | QQQQAPPAPPNPPINTSPINSTLNLSLGTSGGGVG/GV/GTGNPQGNTNTTG | 735 | Db | 644 | -----GPALTSQIN-LANLQARDNA | |
| Oy | 736 | SPSPPOOOQDGTINKSPPIPDARPADALYGDQDITITMPLAYTTSRSHSN | 795 | Oy | 736 | SPSPPOOOQDGTINKSPPIPDARPADALYGDQDITITMPLAYTTSRSHSN | 795 |
| Db | 679 | APRISO-----PRGDALYGDQDITITMPLAYTTSRSHSN | 720 | Db | 679 | APRISO-----PRGDALYGDQDITITMPLAYTTSRSHSN | 720 |
| Oy | 796 | YLAQLOVESFSNNNKITEMPSKITLIGTMPLAUNLVGLGRNVVUPCSEGTTNLGWPPGI | 855 | Oy | 796 | YLAQLOVESFSNNNKITEMPSKITLIGTMPLAUNLVGLGRNVVUPCSEGTTNLGWPPGI | 855 |
| Db | 721 | YLAQLOVESFSNNNKITEMPSKITLIGTMPLAUNLVGLGRNVVUPCSEGTTNLGWPPGI | 779 | Db | 721 | YLAQLOVESFSNNNKITEMPSKITLIGTMPLAUNLVGLGRNVVUPCSEGTTNLGWPPGI | 779 |
| Oy | 856 | VGVLIPIIISFSITPAQWKLKNNI---EGQNQNRIFIN---HSSRSIRSISLGLKK | 907 | Oy | 856 | VGVLIPIIISFSITPAQWKLKNNI---EGQNQNRIFIN---HSSRSIRSISLGLKK | 907 |
| Db | 780 | LGVLILLAVGMPASLYWIKRDPATLNEAASCAKVSISSFLPKRNKRPNDSKINNV | 839 | Db | 780 | LGVLILLAVGMPASLYWIKRDPATLNEAASCAKVSISSFLPKRNKRPNDSKINNV | 839 |
| Oy | 908 | HGG--NKSISIPEPNKY | 921 | Oy | 908 | HGG--NKSISIPEPNKY | 921 |
| Db | 840 | RAGPNKNSVSLFSRY | 855 | Db | 840 | RAGPNKNSVSLFSRY | 855 |
| RESULT 3 | | | | | | | |
| Oy | US-08-945-749-1 | Sequence 1, Application US/08945749 | | Oy | US-08-945-749-2 | Sequence 2, Application US/08945749 | |
| | Publication No. US2020138880A1 | | | | | Publication No. US2003138880A1 | |
| | GENERAL INFORMATION: | | | | | GENERAL INFORMATION: | |
| Oy | 1 | APPLICANT: GARDNER, Richard C | | Oy | 1 | APPLICANT: GARDNER, Richard C | |
| Oy | 2 | APPLICANT: MACDIARMID, Colin W | | Oy | 2 | APPLICANT: MACDIARMID, Colin W | |
| Oy | 3 | APPLICANT: HAY, Robert J | | Oy | 3 | APPLICANT: HAY, Robert J | |
| Oy | 4 | APPLICANT: Auckland Uniservices Limited | | Oy | 4 | APPLICANT: Auckland Uniservices Limited | |
| Oy | 5 | APPLICANT: New Zealand Pastoral Agriculture Research Institut | | Oy | 5 | APPLICANT: New Zealand Pastoral Agriculture Research Institut | |
| Oy | 6 | FILE REFERENCE: 08/945-749 | | Oy | 6 | FILE REFERENCE: 08/945-749 | |
| Oy | 7 | CURRENT APPLICATION NUMBER: US/08/945-749 | | Oy | 7 | CURRENT APPLICATION NUMBER: US/08/945-749 | |
| Oy | 8 | CURRENT FILING DATE: 1998-01-12 | | Oy | 8 | CURRENT FILING DATE: 1998-01-12 | |
| Oy | 9 | EARLIER APPLICATION NUMBER: PCV/NZ96/0035 | | Oy | 9 | EARLIER APPLICATION NUMBER: PCV/NZ96/0035 | |
| Oy | 10 | EARLIER FILING DATE: 1996-01-01 | | Oy | 10 | EARLIER FILING DATE: 1996-01-01 | |
| Oy | 11 | EARLIER FILING DATE: 1995-05-01 | | Oy | 11 | EARLIER FILING DATE: 1995-05-01 | |
| Oy | 12 | NUMBER OF SEQ ID NOS: 16 | | Oy | 12 | NUMBER OF SEQ ID NOS: 16 | |
| Oy | 13 | LENGTH: 859 | | Oy | 13 | LENGTH: 859 | |
| Oy | 14 | TYPE: PRP | | Oy | 14 | TYPE: PRP | |
| Oy | 15 | ORGANISM: Saccharomyces cerevisiae | | Oy | 15 | ORGANISM: Saccharomyces cerevisiae | |
| Oy | 16 | SEQ ID NO: 1 | | Oy | 16 | SEQ ID NO: 1 | |
| Oy | 17 | LENGTH: 859 | | Oy | 17 | LENGTH: 859 | |
| Oy | 18 | Query Match 34.8%; Score 1679; DB 2; Length 859; | | Oy | 18 | Query Match 34.8%; Score 1666; DB 2; Length 858; | |
| Oy | 19 | Best Local Similarity 41.5%; Pred. No. 3; se-107; Mismatches 405; Conservative 137; Indels 176; Gaps 31; | | Oy | 19 | Best Local Similarity 40.9%; Pred. No. 2; se-106; Mismatches 401; Conservative 126; Indels 233; Gaps 35; | |
| Oy | 20 | MSDBSYXNTNQNPISPRDEVLDDEHNRNQITCAISDSEDELE--KSEBLSEYVK 57 | | Oy | 20 | MSDBSYXNTNQNPISPRDEVLDDEHNRNQITCAISDSEDELE--KSEBLSEYVK 57 | |
| Oy | 21 | 1 MSSSS--SSESSPYLSRS-----NLSLANTYNSKTEDTGTYLDIRHDPDLSPRH 49 | | Oy | 21 | 1 MSSSS--SSESSPYLSRS-----NLSLANTYNSKTEDTGTYLDIRHDPDLSPRH 49 | |
| Oy | 22 | 58 EKQOHHQETSTDNAKPLTRKSSSTIKKS----NLTDKDRTIPMPLSUGGDDTNSCH 112 | | Oy | 22 | 58 EKQOHHQETSTDNAKPLTRKSSSTIKKS----NLTDKDRTIPMPLSUGGDDTNSCH 112 | |
| Oy | 23 | 50 QPPLKNUKEAKSTKPKPEQNSATRNSHVDSVGRGMPEDGQMDDEV-AHH 108 | | Oy | 23 | 50 QPPLKNUKEAKSTKPKPEQNSATRNSHVDSVGRGMPEDGQMDDEV-AHH 108 | |
| Oy | 24 | 113 KNR----NYNSSL----RKOFILKDNTDNSTNNT--HALIPPIPITPI 156 | | Oy | 24 | 113 KNR----NYNSSL----RKOFILKDNTDNSTNNT--HALIPPIPITPI 156 | |
| Oy | 25 | 109 QLRSALTSWARSFRLAHSMPLQRQLYBSNI----HTPPPKVGDYMTMSSTA 161 | | Oy | 25 | 109 QLRSALTSWARSFRLAHSMPLQRQLYBSNI----HTPPPKVGDYMTMSSTA 161 | |
| Oy | 26 | 157 TNANKSRKSKQSENLENPLTIKKIGANNRNPFNDLVEPMTCMKNDSEDTNT----210 | | Oy | 26 | 157 TNANKSRKSKQSENLENPLTIKKIGANNRNPFNDLVEPMTCMKNDSEDTNT----210 | |
| Oy | 27 | 162 SSGNKS-KLSASSSSASITKVR-----KSSVSVPLEPKHSDTHSKLAKERK 210 | | Oy | 27 | 162 SSGNKS-KLSASSSSASITKVR-----KSSVSVPLEPKHSDTHSKLAKERK 210 | |
| Oy | 28 | 211 ---STTANHKLKGATTLGCGCTATA---TATAAGGRPSRSITSEADSHASRSQ 264 | | Oy | 28 | 211 ---STTANHKLKGATTLGCGCTATA---TATAAGGRPSRSITSEADSHASRSQ 264 | |
| Oy | 29 | 211 RTYTTDAHSSSIN-PAVLTKSTSQSKSDADDLTKRKPYRNTAFPSDY-SQASRSQ 268 | | Oy | 29 | 211 RTYTTDAHSSSIN-PAVLTKSTSQSKSDADDLTKRKPYRNTAFPSDY-SQASRSQ 268 | |
| Oy | 30 | 265 ETEEDVCFPAYGD-HIRVNCGIDDEDEFIREEEBAYVLRQCMIAK-NILRIDEFQNLISK 322 | | Oy | 30 | 265 ETEEDVCFPAYGD-HIRVNCGIDDEDEFIREEEBAYVLRQCMIAK-NILRIDEFQNLISK 322 | |

; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus
 ; TITLE OF INVENTION: Methods of Use
 ; FILE REFERENCE: 10182-018-999
 ; CURRENT APPLICATION NUMBER: US 10/128,714
 ; CURRENT FILING DATE: 2002-04-03
 ; PRIORITY NUMBER: US 60/285,697
 ; PRIOR FILING DATE: 2001-04-21
 ; PRIORITY NUMBER: US 60/287,066
 ; PRIOR FILING DATE: 2001-04-27
 ; PRIORITY NUMBER: US 60/295,890
 ; PRIOR FILING DATE: 2001-06-05
 ; PRIORITY NUMBER: US 60/303,899
 ; PRIOR FILING DATE: 2001-07-09
 ; PRIORITY NUMBER: US 60/316,362
 ; PRIOR FILING DATE: 2001-08-31
 ; NUMBER OF SEQ ID NOS: 8603
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO: 3563
 ; TYPE: PRT
 ; ORGANISM: Aspergillus fumigatus
 ; US-10-128-714-3563

Query Match 22.8%: Score 1099; DB 4; Length 663;
 Best Local Similarity 32.7%; Pred. No. 3-67;
 Matches 292; Conservative 98; Mismatches 241; Indels 262; Gaps 22;

9 DSSSDLPRSKVDTAAASMKYTKPLKENTRQ---YSDAQPIRHE---ALAKVDETRDS 62
 94 RTINPMISLGQDITNSKHKNRYN---NMS-SURKDIFYKDKTDNSTNTHLAIPIP 148
 63 R--HRPSSNGE--NSGVENGVVEKTNTSTSGRDF--EGEAEEAVKRYOLSPAL- 114
 149 IPIPPIPITNAKSR----RKSQLENPLIKK-----TGRANSNFEEND 191
 115 ----LSSNAPSLPLAKSETHOQIRVEESTAPSPLKAELRGHTDPAQTNSRN 168
 192 LYSPTMTRAKTN-----DSEDINTP-----STTAHMKLIGCATTL 226
 169 EASSSARTTSARKSLVSIAFETSASEENGTHTKQAKLKRRTYPSLTHS--SYNPPTL 226
 227 GUVTGTTATAATAG-----RRPSSSDSEBADSHAARSQETEDUCPMPVG- 276
 227 ----LTTAATSKSOKSINGDIDTRRIPPLRDSSRVSHPHSI-QSARDQETEBCDPCMPL 280
 277 DHTRVNGIDPDEIDEFIREEREARYLQKMTNLRIDFQNLSKNTTGAASHR---P 333
 281 LHTRVNGIDPDELEYA-----QISNAERNISLANTHQHSEET 318
 334 YHHSNNRNRKGDDGSSNAAKYTPNKLTKTSRPPTHENSSEAYLKTRQQ 393
 319 YHHTQDGTFTNSATGSS-AALKYPE---ISR---TLEKNCSTNEYKVSNES 368
 394 PPVKYDQDLSTLSSTSFSGSGCQVKFGGARISDGINGS-----SLPD 437
 369 -VREDDKDL-----HDPVTGPKNIEKEGDDSYSRAYTTLQTEQIPS 415
 438 RFLSPHSEETHTADIPSLSPQSITRDLIFRGEATIDAVFTA 497
 416 RPSFPRSEDTETHASDIPSLSPQSITRDLIFRGEATIDAVFTA 475
 498 HPLTAEDTRMQETREKVELPKYXYYFVCPHTPEAKESSEDTLPINIVYIVPHDGLTIPH 557
 476 HPLTAEDIMQETREKVELPKYXYYFVCPHTPEAKESSEDTLPINIVYIVPHDGLTIPH 535
 558 SPISHPANVRVROLRDYDVSAWLCYALIDITDGPAPVHGLEYEADAEVFTA 617
 536 DPESICANVRVROLRDYDVSAWLCYALIDITDGPAPVHGLEYEADAEVFTA 595
 618 RDTDPSMLQRIGESRRKVMTLMLLSGKADWIKMFAKRCQEANSSGYQRO---YNL 674
 596 RQDFAAMQRIGESRRKVMTLMLLSGKADWIKMFAKRCQEANSSGYQRO---YNL 655
 675 QOOQOAPPPNPNIITSPINSTLNLSTGTTGGVGVINGPNPNTNTNTNT 734
 656 QNEQE-----NVHQ-----NNSN----- 668
 735 GSPSPPQQQHQGITNKSPPIDPAPRADIALYLGDOIHDITPMLONLAYKTPSRHS 794
 669 -----NOISLUSN-SYMOITSQPRGDIALYLGDOIHDITPMLONLAYKTPSRHS 718
 795 NYLAQLOVESFNNSNNKITEMPSKTKLIGTMVLPUVNTGSPGRNVRVPQBGCTNLWFFG 854
 719 NYLAQLOVESFNNSNNKITEMLGKUTMAGTMVLPUVNTGSPGRN-VTAWYQ 777
 855 TCVGLPILISSEPAQWMLKLKLNSTI---EGQGNRNP--- 900
 778 IGVLLVLLAVSWLASTWIKIDDPATLNEANGSAKSVISSPLBKDRKFNDNSKCN 837
 901 RSLGLKKGKNSLISPPNKY 921
 838 ARVGVR-----KSTVSLPSRY 854

Result 4
 US-10-128-714-3563
 Sequence 3463, Application US/10128714
 Publication No. US20030112013A1
 GENERAL INFORMATION:
 ; APPLICANT: Jiang, Bo
 ; APPLICANT: Hu, Wengi
 ; APPLICANT: Tishkoff, Daniel
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Broshkin, Alexey M
 ; APPLICANT: Lemieux, Sebastien M

| | | | |
|--------------------------------|---|-----------------------------|--|
| Qy | 761 RADIALYLGIDODHITMFMQNLAYKEFSSRSNHLAQLOVESFNSNNKITEMPSKITLE 820 | Db | 305 -----; IVERLONENBETRAGFPSSQSSTVHAELGDLVLRGD 342 |
| Db | 548 RDGDIGYLGDIDQHVTMNSLAIFERKLNSHTWILQLAVNTLVUGNNAVKLSKVTL 607 | Qy | 464 SVRDPLRNGE - TWNLDCTCPDSEMKNLAKAFGHPITADIDIMOTREKVEVLFPSKTY 521 |
| Qy | 821 IGMILVPLNLTVLGEGMAMVRVPGEGTGNLGFPGVGLIITIGSTPIQWW 873 | Db | 343 FTRDOLQBGEGSVWVLDFULNPTEEVAAALSFRSHPTTDIDUOAREKVEPKQY 402 |
| Db | 608 IATMLVPMNLJCGLGFMNMVRVPGECQECGLNFPGIVGVIAVITLSSIAARRY 660 | Qy | 522 FVCFHFT-EADKESEDLEPBPINYYTVVFIDGLTHFSPISHANVRVRVRLRDYDVDS 580 |
| RESULT 5 US-10-128-714-8563 | and Sequence 8563, Application US/1028714 ; Publication No. US20030119013A1 GENERAL INFORMATION: APPLICANT: Jiang, Bo APPLICANT: Hu, Wencui APPLICANT: Nishkoff, Daniel APPLICANT: Zamudio, Carlos APPLICANT: Eroshkin, Alexey M APPLICANT: Lemieux, Sébastien M | Db | 581 ADOLCIALDEITDGAPTHGLEYTAADIEADAVTARTDDESDNLQRIGSERKVTLN 640 |
| Qy | TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus | Db | 463 SDWICATMIDDIVSFDPVNPYMFEDGYLSISPPENPAANVRKRGDRYVSLS 462 |
| Qy | TITLE OF INVENTION: Methods of Use | Db | 522 RLUKGADVTKMPAKFCOCBANEASSSYYQRQYLNQQQQQAOPPPNPDTSPINSTNLN 700 |
| Qy | FILE REFERENCE: 1018-018-999 | Db | 548 RGDGILGDBIQDHVMMNLAHFBRMLSRHTNLYLQAVNTLNVLGHNVNKLVKVTL 607 |
| Qy | CURRENT APPLICATION NUMBER: US/10/128-714 | Db | 523 RLUKGADVTRGFSKRCNE ----- 541 |
| Qy | CURRENT FILING DATE: 2002-04-23 | Qy | 701 NSLGTSRGSGGVGGINFENPIGNTNTNSPSP00000HTNKSPIPDARP 760 |
| Qy | PRIOR APPLICATION NUMBER: US 60/285, 697 | Db | 542 ----- OYST----- P 547 |
| Qy | PRIOR FILING DATE: 2001-04-23 | Db | 761 RADIALYLDIDODHITMFMQNLAYKEFSSRSNHLAQLOVESFNSNNKITEMPSKITLE 820 |
| Qy | PRIOR APPLICATION NUMBER: US 60/287, 066 | Db | 548 RGDGILGDBIQDHVMMNLAHFBRMLSRHTNLYLQAVNTLNVLGHNVNKLVKVTL 607 |
| Qy | PRIOR FILING DATE: 2001-04-27 | Qy | 821 IGMILVPLNLTVLGEGMAMVRVPGEGTGNLGFPGVGLIITIGSTPIQWW 873 |
| Qy | PRIOR APPLICATION NUMBER: US 60/295, 890 | Db | 608 IATMLVPMNLJCGLGFMNMVRVPGECQECGLNFPGIVGVIAVITLSSIAARRY 660 |
| Qy | PRIOR FILING DATE: 2001-07-09 | RESULT 6 US-10-945-749-3 | ----- |
| Qy | PRIOR APPLICATION NUMBER: US 60/303, 899 | Qy | ; Sequence 3, Application US/08945749 |
| Qy | PRIOR FILING DATE: 2001-07-09 | Db | ; Publication No. US20020138880A1 |
| Qy | PRIOR FILING DATE: 2001-07-09 | Qy | ; GENERAL INFORMATION: |
| Qy | PRIOR FILING DATE: 2001-08-31 | Db | ; APPLICANT: GARDNER, Richard C |
| Qy | NUMBER OF SEQ ID NOS: 8603 | Qy | ; APPLICANT: MACDARMID, Colin W |
| Qy | SOFTWARE: PatentIn version 3.1 | Db | ; APPLICANT: HAY, Robert J |
| Qy | SEQ ID NO: 8563 | Qy | ; APPLICANT: Auckland UniServices Limited |
| Qy | LENGTH: 663 | Db | ; APPLICANT: New Zealand Pastoral Agriculture Research Institut |
| Qy | TYPE: PRT | Qy | ; TITLE OF INVENTION: ALUMINUM RESISTANCE GENE |
| Qy | ORGANISM: Aspergillus fumigatus | Db | ; FILE REFERENCE: 08/945-749 |
| Qy | US-10-128-714-8563 | Qy | ; CURRENT APPLICATION NUMBER: US/08/945-749 |
| Query Match | Best Local Similarity 22.8%; Score 1099; DB 4; Length 663; | Db | ; CURRENT FILING DATE: 1998-01-12 |
| Qy | Matches 292; Conservative 98%; Mismatches 241; Indels 262; Gaps 22; | Qy | ; EARLIER FILING NUMBER: PCT/NZ96/0035 |
| Qy | 4 SESYYQNSTMTHOPTSDEVDDHNRNQINTNCAFSDSEDEELKSELESEVKSEOOH 63 | Db | ; EARLIER FILING DATE: 1996-05-01 |
| Db | 7 SEKYACLTFSTPVE----DDDRFQI----DSSPRIATEADI-SLSRONTQHAY 54 | Qy | ; EARLIER FILING NUMBER: NZ 272039 |
| Qy | 64 HOBTSDNAKPLTRKSGSIKKSNLTDKDRITNEMSLSGDDTNTSGHKNRNYYNSSSLR 123 | Db | ; NUMBER OF SEQ ID NOS: 16 |
| Db | 55 HOETPQ-----REPDLSSTDQDALRAGSLSRPEQALID 87 | Qy | ; SOFTWARE: PatentIn Ver. 2.0 |
| Qy | 124 KOPYIKD---NTDDNSTH---THAIIPIPPIPITNANKSRK---SQ 167 | Db | ; SEQ ID NO: 3 |
| Db | 88 DRSRGKDNTLGRFSTDPNPNVRHGTTWSRTHOEL-----ANMRESSPSA 134 | Qy | ; LENGTH: 969 |
| Qy | 168 LENPLPLKKTIGRANNNNPFNDLVSPTMKYNDSEDTNTSTTANMKLGIGATTG 227 | Db | ; TYPE: PRT |
| Db | 135 RSSSP-----N5VEAFDRRERANTLE---SHAAPDLEALQTVSG 176 | Qy | ; ORGANISM: Saccharomyces cerevisiae |
| Qy | 228 VGTGT-TATATATAAGRPRSSSIDSEADSHARSQETEVDCPMVGDHTRVNGIDE 286 | Db | US-08/945-749-3 |
| Db | 177 ---GTHPRPTPSWAIRPQDGIOLEPN-----DSCSYTPEQPERIVTNT 222 | Qy | Query Match 17.8%; Score 856; DB 2; Length 969; |
| Qy | 287 DEIDPTEBEERAYOKMAMKILBIDEQNIKNTNSGASPH--XHHNSNTRK 343 | Db | Best Local Similarity 24.1%; Pred. No. 46-51; Gaps 29; |
| Db | 223 BELEEFVA-----LSRQKMPSTSRS-----KQSLSSRGPYFYDYLPLRKL 266 | Qy | Matches 270; Conservative 143; Mismatches 248; Indels 460; |
| Qy | 344 NGGGGGGSMAALKTPRNKLTLSRFETHNNSSETEYSPKQPYKNDQLS 403 | Db | 11 SITNNOPIRSDEVDDHNRNQINTNCAFSDSEDEELKSELESEVKSEOOH----- 58 |
| Qy | 267 DVEGEKRSAS-----DRSSDLMMDAQLTADKVVANYDDERD 304 | Qy | 44 SDSRPQOLLHDNLQHNRQITDPQI-DSGMLRHSUDSTSDNIKSESDPSLKGAFIDHR 102 |
| Db | 404 LTSSTSTSGSGQVKFGARIDQINGSLPDRS1P1SESSETHARDIPSLSWPGO 463 | Qy | 59 ---KQHQHQBETSD-NAKPDRTRKSGSIKKSNLTDKDRITNEMSLSGDDTNTSGHKNRNYYNSSSLR 160 |
| Qy | 226 LGVGRCTTATATAAGRPRSSSIDSEADSHARSQETEVDCPMVGDHTRVNGIDE 283 | Db | 103 PSMSOPREPOSVSTVOQPMKSTPSYKPGLSDPSDQNSRSVLSLSE-LWSL 160 |
| Qy | 216 -----KRR---GSDDESN-KN---SKSSTSDPSDNEBDRNSRKR 215 | Qy | 113 KRNBNMMSLREDFYLNK-TDNSTNNTHLAIPIRIPPLITNA----NKSRRK 165 |
| Qy | 216 -----KRR---GSDDESN-KN---SKSSTSDPSDNEBDRNSRKR 215 | Db | 161 KRR---KSVHKSFDENSPTDRRSNANNDVIVD-ALANHYNNASTGUNDSRKR 215 |

US-09-864-408A-2266

Query Match Score 345.5; DB 3; Length 154;
Best Local Similarity 7.28; Pred. No. 6; 8e-16;
Matches 79; Conservative 20; Mismatches 39; Indels 39; Gaps 6;

Qy 331 GSSMAKTKPKNLKLTKTLPRTNSSESPRTRPLTKQQPKYDQSLTSSTS 410
Db 1 GSS-AALKTPPE----IR---TLEKNSVNTMYSNNNS--VREDKDPL----- 42

Qy 411 TSGSGCCGKFGGARISDGNGG-----SLPDRFLSPHSESETHADP 454
Db 43 455 IPSLVSQPSVYRDLPRGETAWLDCTCPDTSERMLAKGTHPLTAEDIMQTR 511
Qy 98 IPSLISEGQTYELFKCGDPWWLDCSCPDTDEMRCIAKPTGHPPLTAEDIMQTR 154

RESULT 8
US-10-425-115-321932
; Sequence 31312, Application US/10425115
; Publication No. US20100421427A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-2153221.B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 359326
; SEQ ID NO: 321932
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_56667C.1.pep
US-10-425-115-321932

Query Match Score 298; DB 4; Length 94;
Best Local Similarity 6.28; Pred. No. 6.7e-13;
Matches 54; Conservative 15; Mismatches 18; Indels 0; Gaps 0;

Qy 573 LRDDYDVSDMVLCAVLDTESTDGPVAVGIEYADEADAYTARDTDPSMMLORGES 632
Db 1 LRDDYDVSDMVLCAVLDTESTDGPVAVGIEYADEADAYTARDTDPSMMLORGES 632

Qy 633 RKKYTMELRLSKGADYKVKFAKCQE 660
Db 61 RKKYTMELRLSKGADYKVKFAKCQE 88

RESULT 9
US-08-945-749-8
; Sequence 8, Application US/08945749
; Publication No. US200802013888A1
; GENERAL INFORMATION:
; APPLICANT: GARDNER, Richard C
; APPLICANT: MACLARND, Colin W
; APPLICANT: HAY, Robert J
; APPLICANT: Auckland Universires Limited
; TITLE OF INVENTION: NEW ZEALAND PASTORAL AGRICULTURE RESEARCH INSTITUT
; FILE REFERENCE: 08/945,749
; CURRENT APPLICATION NUMBER: US/08/945,749
; CURRENT FILING DATE: 1998-01-12
; EARLIER APPLICATION NUMBER: PCT/NZ96/0035
; EARLIER FILING DATE: 1996-05-01
; EARLIER APPLICATION NUMBER: NZ 272039
; EARLIER FILING DATE: 1995-05-01
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO: 8
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Synechocystis sp.

US-08-945-749-8

Query Match 6.2%; Score 298; DB 2; Length 387;
 Best Local Similarity 23.0%; Pred. No. 4.8e-12;
 Matches 108; Conservative 70; Mismatches 143; Index 148; Gaps 18;

Oy 437 DRFSLPH---SESEETIHADIPS---LV---SPGOSVR---DLEFNGBETW----W 477
 Db 27 DFDYDDEPGSEPOTSLIEPAPSRIVLVDPSHARVKSDISPALRPLGINTVSW 86

Oy 478 LDCT---CPTDSKEMKLMARAFGHITPQHEDIRQETRKEVLFKYTYVCPHTPADEKED 536
 Db 87 MDIEGLGSEEVKEKEYGEIJKHPULLSDIVNPRAKVEDYNDHVIAHYRPNRE-BD 145

Oy 537 YLEPINV---YIVVPHOCILTEPSPISHANVRVRVQLRDYU-DVSADMILCYXL 588
 Db 146 GPSEGVSPVSKYRLTQPSEHLIDCNP---REFIRTGKRVCOQGADLYXL 199

Oy 589 IDEITDGPAVIGHGLEYEADAEVARTADFTESMQLORIGBSRKVATLRLSGRAD 648
 Db 200 IDMLIDEVPLLEDYEEALEDTI---RNPN-SSMLELYHIRELLALARLI---- 251

Oy 649 VIKMPAKRQCBANSSGTYQRQNLQOOOOQAPPNPMLITSPINSTNLNSLGTSNG 708
 Db 252 WPLRYMANWLDRD--- 265

Oy 709 GSVGVGINFGENPGNTNTNTGSPSPPOOOGHQITNKSFIDARPRAIDLXL 768
 Db 266 -----ADRVTRP 278

Oy 769 GDQDHITMFONLAYERKISRSHSNIALQLOVESNSNKATPSKIMTIGMALYPL 828
 Db 279 RDCYDHITQVLDITIAYRELASLMEYMTAM---SNKONETMKFLUTISTIFITPL 331

Oy 829 NLVYGLFGHNRVWPGEGTNLUWGPFGIVGULIPIIIGSFIACOW---WL 874
 Db 332 TPIAGYGRANTFENPE---LNSRGYITIWMLIAGGSLYFFWRGWL 378

RESULT 10

Sequence 7005, Application US/11097143

; GENERAL INFORMATION:

; APPLICANT: Venter, J. Craig

; ET AL.

; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEARIC ACID

; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE

; FILE REFERENCE: CL000728

; CURRENT APPLICATION NUMBER: US/11/097.143

; CURRENT FILING DATE: 2005-04-04

; PRIOR APPLICATION NUMBER: 60/157,832

; PRIOR FILING DATE: 1999-10-05

; PRIOR APPLICATION NUMBER: 60/160,191

; PRIOR FILING DATE: 1999-10-19

; PRIOR APPLICATION NUMBER: 60/161,932

; PRIOR FILING DATE: 1999-10-28

; PRIOR APPLICATION NUMBER: 60/164,789

; PRIOR FILING DATE: 1999-11-12

; PRIOR APPLICATION NUMBER: 60/173,383

; PRIOR FILING DATE: 1999-12-28

; PRIOR APPLICATION NUMBER: 60/175,693

; PRIOR FILING DATE: 2000-01-12

; PRIOR APPLICATION NUMBER: 60/184,831

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: 60/191,637

; PRIOR FILING DATE: 2000-03-23

; NUMBER OF SEQ ID NOS: 43008

; SOFTWARE: PastSeq for Windows Version 4.0

; SEQ ID NO: 7905

; LENGTH: 1198

; TYPE: PRO

; ORGANISM: DROSOPHILA

; US-11-097-143-7905

; Query Match 4.9%; Score 234.5; DB 6; Length 1198;

; Best Local Similarity 19.7%; Pred. No. 5.7e-07;

; Matches 114; Mismatches 291; Index 349; Gaps 41;

Qy 57 SEKOQQHHQEITSNAKPLTRKGSS-----IKKS----- 87
 Db 37 SHQQQQQQOLMPHHKDQMIAQASSPMLPYSHLQLQEDATAITGPAAAAYEAATT 96
 Qy 88 -----WLTQDKDRTINPMPSLSCGDDR---INSCHKVNOMN---SSLRKDFYLKDN 131
 Db 97 SANADNFSSLLTIDASQDGGISLUGCDRPPVAPSPHNSNMNTLMGTATATTTTN 156
 Qy 132 TDDNSTNTHLAIPIPPIP-----PLTNAKRSQSKQBNLPLPKKT 179
 Db 157 NNNNNNNNNNNNNEAKTVPSNSVIESVTPMSFANILPPTRHSANECDPDLQZN 216
 Qy 180 IGRNNNNNPFENDLVP-----MPKMKTNDEDTLN---TSTPAN--HNKLIG--- 222
 Db 217 PONPNGNN---SS11UPPVETQKPLEVNSTSSTSPNLSQSTQDLPFVQGRDGH 275
 Qy 223 -ATPLYGQT-----TATATATAAAGR-RPERSI----- 251
 Db 276 SMTTTRPGCSASGSGSGSGSGSGSTARTGATPTTISNTANTPRTSLSHTELA 335
 Qy 252 -----DSEADSHAHSRQSSQEPEDCPCPMVGDHIRVN-----GIDPD 287
 Db 316 ASSCAPRASPNRHTSASTPOOOOOOHQSGRHSSNSLSDDBMSDDEFLD 395
 Qy 288 E1DEFIREERAYLQKQMIAKNLIRDEFQ-----NLUSKNNTSG-----ASRH 332
 Db 396 DNGEMDQOQATVQISSUNSPFSMOLQEQOOQSSPALAAGNSNNAASGNNNNASGN 455
 Qy 333 PYHHHSNNKKAN-----GDDGSSSSMALKTT----- 360
 Qy 456 NTSSSSNNNNNNNNNDNDAAHV1TKPEHYNAYT-QLGGGSGGSGQHHSNNSNNGNHQ 515
 Qy 361 -----PKNTKLTTSRPEFHENSSESYELKTKQPPY 396
 Db 516 QQQQQQQQQQOHQQQQQEHYQQQQQNMIANA-NQF---NESSYSYIYNPDSQYTPP 569
 Qy 397 KYDDOLSLSTSSTSSTSSGSGQVKFGARIISDGINGS-----LPDRFLPSSES 448
 Db 510 GYQD---TTSHSOQSGGGG---GGNLLNSGSSSGAGCMLP----- 611
 Qy 449 TIHDADIPSLSVSPQOSTRDIFRNGEETWWLDCPCPDSEKMLAKAFG----- 497
 Db 612 -----QANSSGGNN-----PNAIGMSSGSGNGGGAGGN 648
 Qy 438 ---HPLTADJRNQETREV-----UFKSTYFCRHFADRESDYLEPINVYVPHD 550
 Db 619 SGPNPNCMGTSATPSHGEGYIDPKHFLPEBPCV-----GDKVSG-----YX 691
 Qy 551 GILTHPSHSPISHANVRVRYQRLDYDVSADWLCYALIDETDGFAPWTH-----GI 603
 Db 632 QLILCE---SCKGPFRTVQKVKVTCV-AEBSCH---IDKPKRRCPCRFQCKLVEGM 744
 Qy 604 EYADEADAFTARDYDSSMLOR-----IGESRKVMTLRLSGRADVIRKMPKRCQ 658
 Db 745 KLEA-VRADRNGRGERN-KPMPYTRDARLQMRQRLQALRNLNSGPDK----- 795
 Qy 659 EANANSSGGYTORQNLQQQQ-----OAPPDPNPNIITSPINSTLNNSLSTG 708
 Db 756 PTIPTSPQYQAYNPKIKIJOQPVSSLTQPDSPFPI-----AIALQCNWASTS 847
 Qy 769 -----GGVGVGGINFGPNTGNTNTNTNTNTG 736
 Db 848 GVIAPEKNAQGCGGGGLN-GPSVSGNGNSNGSSNGN 885

RESULT 11

US-08-945-749-7

; Sequence 7, Application US/08945749

; Publication No. US2008138880A1

; GENERAL INFORMATION:

; APPLICANT: HAY, Robert J

; APPLICANT: Auckland Uniservices Limited

; APPLICANT: New Zealand Pastoral Agriculture Research Institut

; TITLE OF INVENTION: ALUMINUM RESISTANCE GENE

; FILE REFERENCE: 08/945,749

; CURRENT APPLICATION NUMBER: US/08/945,749

CURRENT FILING DATE: 1998-01-12
 ; EARLIER APPLICATION NUMBER: PCT/NZ96/0035
 ; EARLIER FILING DATE: 1996-05-01
 ; EARLIER APPLICATION NUMBER: NZ 272039
 ; EARLIER FILING DATE: 1995-05-01
 ; NUMBER OF SEQ ID NOs: 16
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO: 7
 ; LENGTH: 380
 ; TYPE: PRT
 ; ORGANISM: Synochocystis sp.
 US-08-945-749-7

Query Match 4.48; Score 214; DB 2; Length 380;
 Best Local Similarity 19.28; Pred. No. 3.1e-06;
 Matches 83; Conservative 72; Mismatches 141; Indels 140; Gaps 14;

Qy 478 LOCCTPSE -----MRLAKARGLTALDRETRKYLEFKSYF 522
 Db 59 LDCEAVDTESVSWINIDGLNHTWRLGVEPLHPVALEDVNQPKVVEYHHL 118
 Qy 523 VCFPHFEADKESEDLEPINVYIVPHDGLTTPHFSF-ISHPAANVRVQLRDYV-DVS 580
 Db 119 PISRNYLDSQQTISEQISPLGKH-YLTTCBESKTYCCLSVERIKTKAGAIRQKN 177
 Qy 581 ADWLYALIDETDTPAVIGIEYEADATEDAYFTARDTDFSSMLQRTGESRKVMTLM 640
 Db 178 ADWLYALIDATIDSPFPV--MEVYCBELQUSUSEIISCPTMKSALKHQODDLMR 234
 Qy 641 RLSKGADVKMPKARCOEANSSSGYQRYNLLQQQQQAPPBPNNPLITSPINSTNL 700
 Db 235 RAIPFORDAI-----244
 Qy 701 NSLGNTSGCGVGGINFGPNTGNTNTGSPSPPQQQQHGNTKSPPIPDARP 760
 Db 245 NSL-----257

RESULT 13
 US-10-156-761-10644
 ; Sequence 10644, Application US/10/156761
 ; Publication No. US2003/115018A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OKURA, SATOSHI
 ; APPLICANT: IKEDA, HARUO
 ; APPLICANT: ISHIKAWA, JUN
 ; APPLICANT: SHIBA, TADAYOSHI
 ; APPLICANT: SAKAI, YOSHIOUKI
 ; APPLICANT: HATORI, MASAHIRA
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-252
 ; CURRENT FILING DATE: 2002-05-30
 ; PRIOR APPLICATION NUMBER: JP 2001-204089
 ; PRIOR FILING DATE: 2001-05-30
 ; PRIOR APPLICATION NUMBER: JP 2001-272697
 ; NUMBER OF SEQ ID NOS: 15109
 ; SEQ ID NO: 10644
 ; LENGTH: 371
 ; TYPE: PRT
 ; ORGANISM: Streptomyces avermitilis
 US-10-156-761-10644

Query Match 4.28; Score 201; DB 4; Length 371;
 Best Local Similarity 19.7; Pred. No. 2.3e-05;
 Matches 91; Conservative 64; Mismatches 160; Indels 148; Gaps 13;

Qy 436 PDRESLPHSSSE-ETTHAPIPSUV-----SPOSUPDLPNCETW 476
 Db 13 PSRSMKDKSADTTDPTPSAVVDCAYTRGARVATDRPLTHEAMQRVRRGGV- 71
 Qy 477 WLDCTCPITDSMKMLAKAFGHPLATEDETRMQTREKVBFLKSYTFVCP--HTEADK- 532
 Db 72 WIGLHEPTEAFAGLAEFLHPLAYEDAVQAHQREKLRYDSDLFTFKTHVHEHQQL 131
 Qy 533 -ESDYLEPINYVYVPHDGLTPHFSFISHPAANVRVQLRDYVDSADWLYALDE 591
 Db 132 TANSEVVETGEVCMCPTRDPFIVTRGGQSLSRALHRQDOPBLJRGPSAVLRAJDH 191
 Qy 592 ITDGPAVIGIEYEADATEDAYFTARDTDFSSMLQRTGESRKVMTLM 645

Query Match 4.48; Score 213.5; DB 4; Length 386;
 Best Local Similarity 19.68; Pred. No. 3.4e-06;

| | | | | | | | |
|---|-----|--|-----|---|-----|---|-------------------|
| Db | 192 | WVDGLAVADAVQDDIDEVEVSPRKRPDKPTDAG---RIVQLKREVLEFRKAVIWPSP | 247 | Db | 239 | -----IIDM----- | -----OHTTSVTP 251 |
| Oy | 646 | KADVIKMPAKRCQFANSSSCKYDQRINLQOCQQCAAPPNNITSPINSTLNNSLGT | 705 | Oy | 740 | FQOORQHITNKSPPIDPARDIAYLGDIODHITMFQNLLAYEKIFSR--SHSY 796 | |
| Db | 248 | LURPQMLSER-----PMRLID----- | 264 | Db | 252 | VVQR-----LAQD-----IRSGSEELFRAYLDQDADHLTRDNTRVSEYRESLSQILNNATL | 304 |
| Oy | 706 | STGGGVGCGTNGFPNGNPNTNTGSPSPQQQQGIGTINNSPPIDPARDIA | 765 | Oy | 797 | LAQIQLQESPNSNNKITEMSKITLIGTMLVPLMLVTCGAMTVRCEGGTNLGHPRG-- | 854 |
| Db | 265 | -----PDIQ 268 | | Db | 305 | VAQRO-----NEDMKRISGWAIIAPFLVSSYGMNDIAPE---LHWAFGTP | 350 |
| Oy | 766 | LYLGDIODHITMFQNLLAYEKIFSRSHSNYLALQLOVESNSNKITEMSKITLIGTML | 825 | Oy | 835 | --IVGVLIIFTGSPITI--AQW 872 | |
| Db | 269 | KYFROYADHLARQBOVLGFIELNLSTQQLANLQAQSAQNEQRKTS-WAATIAVTP- | 326 | Db | 351 | LALLAMIGFTULLWIKRSKWI 372 | |
| Oy | 826 | VPLNLYTGSLGMVNRVCEGGTNLGMWFG--IVGVLIIFTI 85 | | | | | |
| Db | 327 | -----VGVYGRNF----EHPELHWRFQGYPVIMSVVVVLCUG 360 | | | | | |
| RESULT 15 | | | | | | | |
| US-11-084-389-248 | | | | | | | |
| Sequence 248, Application US110842389 | | | | | | | |
| Publication No. US201050244935A1 | | | | | | | |
| GENERAL INFORMATION: | | | | | | | |
| ; Sequence 1568, Application US/09738626 | | | | | | | |
| ; Publication No. US201050244935A1 | | | | | | | |
| ; APPLICANT: NAKAGAWA, SATOSHI | | | | | | | |
| ; APPLICANT: MIZOGUCHI, HIROSHI | | | | | | | |
| ; APPLICANT: ANDO, SEIJI | | | | | | | |
| ; APPLICANT: HAYASHI, MIKIRO | | | | | | | |
| ; APPLICANT: OCHIAI, KEIKO | | | | | | | |
| ; APPLICANT: YOKOI, HARUHIKO | | | | | | | |
| ; APPLICANT: TATEISHI, NAOKO | | | | | | | |
| ; APPLICANT: SENOH, AKIHIRO | | | | | | | |
| ; APPLICANT: IKEDA, MASATO | | | | | | | |
| ; APPLICANT: OZAKI, AKIO | | | | | | | |
| ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES | | | | | | | |
| ; FILE REFERENCE: 249-125 | | | | | | | |
| ; CURRENT APPLICATION NUMBER: US/09/738,626 | | | | | | | |
| ; CURRENT FILING DATE: 2000-12-18 | | | | | | | |
| ; PRIORITY APPLICATION NUMBER: JP 99/377484 | | | | | | | |
| ; PRIOR FILING DATE: 1999-12-16 | | | | | | | |
| ; PRIORITY APPLICATION NUMBER: JP 00/159162 | | | | | | | |
| ; PRIOR FILING DATE: 2000-04-07 | | | | | | | |
| ; PRIORITY APPLICATION NUMBER: JP 00/280988 | | | | | | | |
| ; PRIORITY FILING DATE: 2000-08-03 | | | | | | | |
| ; NUMBER OF SEQ ID NOS: 7059 | | | | | | | |
| ; SOFTWARE: PatentIn ver. 3.0 | | | | | | | |
| ; SEQ ID NO: 3568 | | | | | | | |
| ; LENGTH: 373 | | | | | | | |
| ; TYPE: PRF | | | | | | | |
| ; ORGANISM: Corynebacterium glutamicum | | | | | | | |
| US-09-738-626-3568 | | | | | | | |
| Query Match Score 1.8% Best Local Similarity 19.9% Matches 100: Conservative 69; Mismatches 144; Indels 189; Gaps 20; | | | | | | | |
| Db | | | | | | | |
| Oy | 432 | GGATISDGINGGSLPDRFLSPHSSESETIADPILVSP-----GQSYR 466 | | Query Match Score 3.7% Best Local Similarity 19.9% Matches 100: Conservative 68; Mismatches 145; Indels 189; Gaps 20; | | | |
| Db | 9 | GAIRRDQMLRPRYIPD-----SANSTPPEVSPILVMTDGPKRPFQGAVTA 57 | | Db | 467 | D-LPRNGETFWLDCTCPDSE--NMMLAKAGPHPLTAEDTRQETKVLPKSYTF 522 | |
| Oy | 467 | D-LPRNGETFWLDCTCPDSE--NMMLAKAGPHPLTAEDTRQETKVLPKSYTF 522 | | Db | 58 | DGLKPGASNRMNTSPASPKTAEWDLHPTIVEDLQGQPKDHYDIF 117 | |
| Db | 58 | DGLKPGASNRMNTSPASPKTAEWDLHPTIVEDLQGQPKDHYDIF 117 | | Oy | 523 | VCFTH-----FEADKSESDYLEPTINYTVVFH-----GQSVR 466 | |
| Oy | 523 | VCFTH-----FEADKSESDYLEPTINYTVVFH-----GQSVR 466 | | Db | 118 | IATRARIYDREDFSE-----PHLKPQIAILCDNQDTSASF---SN 166 | |
| Db | 118 | IATRARIYDREDFSE-----PHLKPQIAILCDNQDTSASF---SN 166 | | Oy | 563 | PANVRVRYQD--RDYDVDSAWMLCYALIDEITDGPVINGIEYATEDAYFTARD 619 | |
| Oy | 563 | PANVRVRYQD--RDYDVDSAWMLCYALIDEITDGPVINGIEYATEDAYFTARD 619 | | Db | 167 | PEEDIKRKTTLADEELLSSGPRAVAYLVIDQDQBLERQF---222 | |
| Db | 167 | PEEDIKRKTTLADEELLSSGPRAVAYLVIDQDQBLERQF---222 | | Oy | 620 | TDPSSMLQIGRSRKRKYMILMLSSGKADYKIMFARCOBEANSSSGYQORNQOOQ 679 | |
| Oy | 620 | TDPSSMLQIGRSRKRKYMILMLSSGKADYKIMFARCOBEANSSSGYQORNQOOQ 679 | | Db | 221 | SGDAAYA-----BRYNUISQS-- 238 | |
| Db | 221 | SGDAAYA-----BRYNUISQS-- 238 | | Oy | 563 | PANVRVRYQD--RDYDVDSAWMLCYALIDEITDGPVINGIEYATEDAYFTARD 619 | |
| Oy | 680 | QAPPPPNNPITSPINSTLNNSLGTGSTGVEGGINFGNPNTNTNTGSSP 739 | | Db | 167 | PEEDIKRKTTLADEELLSSGPRAAAYLIDATDQFSPVLRGIAIDQTERQVP-- 222 | |

Qy 620 TDFFSMQLRIGESRKVNTLMLSGKADVIKPAKRCQEANSSGGYTORQYNLQQQQ 679
 Db 223 -----SGDATA-----:|:|:|:
 Qy 680 QAPPPEPNITSPINSLNLSLOTSTGGGGVGGINFGNPNTGSPSP 739
 Db 239 -----ITDM-----:|:|:
 Qy 740 PQQQOHGTTRKSPPIPARDPARDIALYLDIOPHITHMFQNLLAEKPSR---SHSNY 796
 Db 252 WQR----LNKDF--IRSGKSEBIRAYLDDVAHLTRDITRSEYRESQLAVNATL 304
 Qy 797 LAQOLYEVSENNKITEMFSKTILGTMLYPLNLVGLFCRANVRVCEGCTNLGWFFG-- 854
 Db 305 VAQRQ----NEDMKKSIGMAAIIAPLVISSYGMFDIPE---LHNAYGP 350
 Qy 855 --IVGVLIITIGSPIF--AWW 872
 Db 351 LALLAMLGFTLLWIFRSRM 372

↙

Search completed. June 13, 2006, 14:23:59

Job time : 187 secs

| | | | |
|--|--|-----------|-----|
| Adn47899 | Thermococ | | |
| Drosophil | Aeb53775 | Aeb53775 | |
| Biocceleration Ltd. | Aab18199 | Aab18199 | |
| | Abd6529 | Abd6529 | |
| | Aae30116 | Aae30116 | |
| | Aef1370 | Aef1370 | |
| | Aef3379 | Aef3379 | |
| | Aef3399 | Aef3399 | |
| | Abj18184 | Abj18184 | |
| | Akm72734 | Akm72734 | |
| | Abp56676 | Abp56676 | |
| | Abp74039 | Abp74039 | |
| | Adel15647 | Adel15647 | |
| | Abd60403 | Abd60403 | |
| | Abb68595 | Abb68595 | |
| | Abu16000 | Abu16000 | |
| | Aeg81142 | Aeg81142 | |
| | Abd74419 | Abd74419 | |
| | Abu43255 | Abu43255 | |
| | Abp40261 | Abp40261 | |
| | Abd67128 | Abd67128 | |
| | Aab19772 | Aab19772 | |
| | Abp51785 | Abp51785 | |
| | Aay28712 | Aay28712 | |
| | Aeo22556 | Aeo22556 | |
| GenCore version 5.1.9 | US-10-018-105A-4 | ADN47899 | |
| Copyright (c) 1993 - 2006 | US-10-018-105A-4 | ADN47899 | |
| Biocceleration Ltd. | US-10-018-105A-4 | ADN47899 | |
| | US-10-018-105A-4 | ADN47899 | |
| run on: | June 13, 2006, 13:59:19 : Search time 200 Seconds (without alignments) 2107.768 Million cell updates/sec | | |
| title: | US-10-018-105A-4 | | |
| perfect score: | 4820 | | |
| sequence: | 1 MSDESYEQNSTNTQPIPRS.....LGLKHHGGKTSIISPENKYE | 922 | |
| scoring table: | BLOSUM62 | | |
| | Gapext 0.0 , Gapext 0.5 | | |
| searched: | 2599619 seqs., 457216429 residues | | |
| total number of hits satisfying chosen parameters: | 2589679 | | |
| minimum DB seq length: | 0 | | |
| maximum DB seq length: | 2000000000 | | |
| | 44 | 159.5 | 3.3 |
| | 45 | 159.5 | 3.3 |
| | 45 | 159.5 | 3.3 |
| | 45 | 159.5 | 3.3 |

ALIGNMENTS

| SUMMARIES | | | | | | |
|-----------|-------|--------------|------|------------|----------------------|----|
| | Score | Match Length | DB | ID | Description | |
| 1: | 4820 | 100 - 0 | 922 | 3 AAB36515 | Aab36515 Candida a | XX |
| 2: | 4814 | 99 - 9 | 922 | 5 Abp73838 | Abp73838 Candida a | XX |
| 3: | 1679 | 34 - 8 | 859 | 2 AAW07872 | Aaw07872 Aluminium | PR |
| 4: | 1666 | 34 - 6 | 858 | 2 AAW07873 | Aaw07873 Aluminium | XX |
| 5: | 1099 | 22 - 8 | 663 | 6 ABJ25905 | Abj25905 Aspergilll | XX |
| 6: | 1099 | 22 - 8 | 663 | 6 ABJ26505 | Abj26505 Aspergilll | XX |
| 7: | 865 | 17 - 9 | 969 | 7 ADK63466 | Adk63466 Disease t | XX |
| 8: | 821 | 17 - 0 | 228 | 5 Abp05970 | Abp05970 Human ORP | DR |
| 9: | 345.5 | 7.2 | 154 | 5 ABP21160 | Abp21160 Human gly | DR |
| 0: | 319.5 | 6.6 | 107 | 5 Abp06753 | Abp06753 Human ORP | XX |
| 1: | 251 | 5.2 | 326 | 9 ABM52411 | Abm52411 N. xanthu | PT |
| 2: | 234.5 | 4.9 | 1198 | 4 ABB60371 | Abb60371 Drosophil | PT |
| 3: | 210.5 | 4.4 | 354 | 9 ABB40726 | Aeb40726 L. pneumo | PT |
| 4: | 210.5 | 4.4 | 357 | 9 AEB37409 | Aeb37409 L. pneumo | XX |
| 5: | 186.5 | 3.9 | 1072 | 5 ABB54963 | Abbs4963 Lactococc | PS |
| 6: | 182.5 | 3.8 | 373 | 4 AAG9814 | Aag9814 C. glutamini | XX |
| 7: | 179.5 | 3.7 | 373 | 3 AED4914 | Aed4914 Membrane | CC |
| 8: | 179.5 | 3.7 | 1172 | 5 ABG93108 | Abg93108 C. albica | CC |
| 9: | 179 | 3.7 | 1199 | 4 ABB38274 | Abb5274 Drosophil | CC |
| 0: | 175 | 3.6 | 1246 | 4 ABB61547 | Abb61547 Drosophil | CC |

IMMATES

| result | No. | Score | Query | Match | Length | DB | ID | Description |
|--------|-------|-------|----------|-------|--------|----------|----|---------------------|
| 1 | 4820 | 100.0 | AAB66515 | 922 | 3 | AAB66515 | | Aab66515 Candida a |
| 2 | 4814 | 99.9 | ABP73838 | 922 | 5 | ABP73838 | | Abp73838 Candida a |
| 3 | 1679 | 34.8 | AAW07873 | 859 | 2 | AAW07873 | | Aaw07873 Aluminium |
| 4 | 1666 | 34.6 | ABJ25905 | 858 | 2 | ABJ25905 | | Abj25905 Aspergilll |
| 5 | 1099 | 22.8 | ABJ26505 | 663 | 6 | ABJ26505 | | Abj26505 Aspergilll |
| 6 | 1099 | 22.8 | ADK24466 | 663 | 6 | ADK24466 | | Adk24466 Disease t |
| 7 | 865 | 17.9 | ADK24466 | 969 | 7 | ADK24466 | | Adk24466 Disease t |
| 8 | 821 | 17.0 | ABP05970 | 228 | 5 | ABP05970 | | Abp05970 Human ORF |
| 9 | 345.5 | 7.2 | ABP21160 | 154 | 5 | ABP21160 | | Abp21160 Human gly |
| 10 | 319.5 | 6.6 | ABP06753 | 107 | 5 | ABP06753 | | Abp06753 Human ORF |
| 11 | 251 | 5.2 | ABM22441 | 326 | 9 | ABM22441 | | Abm22441 M. xanth |
| 12 | 234.5 | 4.9 | ABB0371 | 1198 | 4 | ABB0371 | | Abb0371 Drosophil |
| 13 | 210.5 | 4.4 | ABB04726 | 354 | 9 | ABB04726 | | Abb04726 L. pneumo |
| 14 | 210.5 | 4.4 | ABB04740 | 357 | 9 | ABB04740 | | Abb04740 L. pneumo |
| 15 | 186.5 | 3.9 | ABB04963 | 1072 | 5 | ABB04963 | | Abb04963 Lactococc |
| 16 | 182.5 | 3.8 | AAGS9814 | 373 | 4 | AAGS9814 | | Aag89814 C glutami |
| 17 | 179.5 | 3.7 | ABG33309 | 373 | 9 | ABG33309 | | Abg33309 C. albica |
| 18 | 179 | 3.7 | ABG33309 | 1172 | 5 | ABG33309 | | Abg33309 C. albica |
| 19 | 175 | 3.7 | ABBS274 | 1199 | 4 | ABBS274 | | Abbs274 Drosophil |
| 20 | 175 | 3.6 | ABR61547 | 1246 | 4 | ABR61547 | | Abbr61547 Drosophil |

useful as drug targets. Fragments of them are useful as probes and primers for diagnosis of fungal infections, also as antisense and ribozyme agents. Proteins encoded by the genes are used to screen for their specific inhibitors which are potential antifungal agents for controlling a wide range of fungi pathogenic on animals or plants.

Fragments of the proteins are also used to raise specific antibodies.

Higher animals do not contain genes closely similar to the C. albicans genes, so antifungal agents that target them are less likely to be toxic to humans.

XX

SQ Sequence 922 AA:

| Query Match | Score | DB | Length |
|-----------------------|--------|--------------|--------|
| Best Local Similarity | 100.0% | 4820; | 922; |
| Matches | 922; | Pred. No. 0; | |
| Conservative | 0; | Mismatches | 0; |
| Indels | 0; | Gaps | 0; |

Oy

1 MSDESYNTSTNTPRSPDEVLDHHRNQTMCATSDSEBLKSELESEVRSKQ 60
 ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 1 MSDESYNTSTNTPRSPDEVLDHHRNQTMCATSDSEBLKSELESEVRSKQ 60

Db

61 QHHQEITSDNAAKPLTRKGSSKKSMNDKDRITNPMLSGDDPDTNGKHNANMNS 120
 ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 61 QHHQEITSDNAAKPLTRKGSSKKSMNDKDRITNPMLSGDDPDTNGKHNANMNS 120

Oy

121 SLRKDFYKLKDNTDSTNNTHLAIPIPPIPPIPITANKSRSKQSLNPLPKTKT 180
 ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 121 SLRKDFYKLKDNTDSTNNTHLAIPIPPIPPIPITANKSRSKQSLNPLPKTKT 180

Oy

181 GRNNSNFENDLVSPATRKTNDTSITNSTTANHMKLGIGATTGVRGGTTATAVTA 240
 ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 181 GRNNSNFENDLVSPATRKTNDTSITNSTTANHMKLGIGATTGVRGGTTATAVTA 240

Db

241 AAGRPRPSRTSIDAHDHSARSQSTEIDVCPMPVGDHIVNGIDFDDEDEFIREBREA 300
 ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 241 AAGRPRPSRTSIDAHDHSARSQSTEIDVCPMPVGDHIVNGIDFDDEDEFIREBREA 300

Oy

301 YLQKQMTAKNKLIRIDEFNLSKKNTTSGASRHPYHISNNNNKNNNGDGSSMSMALYKT 360
 ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 301 YLQKQMTAKNKLIRIDEFNLSKKNTTSGASRHPYHISNNNNKNNNGDGSSMSMALYKT 360

Oy

361 PRNIUKLTLSPEFHENSSESETIYLTKTQQPKYDQSLSTSTSTSGSGSVYK 420
 ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 361 PRNIUKLTLSPEFHENSSESETIYLTKTQQPKYDQSLSTSTSTSGSGSVYK 420

Db

421 FGGSARISDINGSLPDRPLSPHSESESETIHAPIPSLVSQGSDVRLFRNGETWMDC 480
 ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 421 FGGSARISDINGSLPDRPLSPHSESESETIHAPIPSLVSQGSDVRLFRNGETWMDC 480

Oy

481 TCTPTDSEKMKAFCIHLPAEDIRMQETREKVELFKSTYFCFPTFADEKESDYLEP 540
 ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 481 TCTPTDSEKMKAFCIHLPAEDIRMQETREKVELFKSTYFCFPTFADEKESDYLEP 540

Db

541 INYVYVPHDGLTTHSPKSHPANVRYTQLRDVDSADMCLTYALDETDGPAPV1 600
 ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 541 INYVYVPHDGLTTHSPKSHPANVRYTQLRDVDSADMCLTYALDETDGPAPV1 600

Oy

601 RGEYEAIDAIDBAIDVAPTARDTDSSENLRIGSRKKTWLMRLLSGADYVTKPARKCQB 660
 ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 601 RGEYEAIDAIDBAIDVAPTARDTDSSENLRIGSRKKTWLMRLLSGADYVTKPARKCQB 660

Db

661 ANSSGGYQROYNLQOQQQAPPPPNPITTSPINSTLNLSIGTSTGGVGVGG1NFGP 720
 ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 661 ANSSGGYQROYNLQOQQQAPPPPNPITTSPINSTLNLSIGTSTGGVGVGG1NFGP 720

Oy

721 NPGANTNTNTGSPSPQQQQHGTINKSPPIPDPARADIALYLDIODHITMfq 780
 ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 721 NPGANTNTNTGSPSPQQQQHGTINKSPPIPDPARADIALYLDIODHITMfq 780

Db

781 NLLAYEKIFRSHSNYLAQLOVESFSNNNKITEMPSKIKLIGTMVPLNLVTGLFGMVR 840
 ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db

781 NLLAYEKIFRSHSNYLAQLOVESFSNNNKITEMPSKIKLIGTMVPLNLVTGLFGMVR 840
 Oy

841 VPGEGGTGTLGMFFGIVGVLFITIIGSPFIAQWMLKLNNSIECONGNRPINPHSSRSSI 900
 ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db

841 VPGEGGTGTLGMFFGIVGVLFITIIGSPFIAQWMLKLNNSIECONGNRPINPHSSRSSI 900
 ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Oy

901 RSLGUKKHGKNSKTSISPNEYE 922
 ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db

901 RSLGUKKHGKNSKTSISPNEYE 922
 ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 2

ABP73838

ID ABP73838 standard; protein: 922 AA.

XX

AC ABP73838;

XX

DT 30-JAN-2003 (first entry)

XX

DE Candida albicans essential protein SEQ ID NO 7675.

XX

KW Fungus; yeast; tetracyclin; promoter; GRACE strain; biosynthesis;

KW signal transduction; DNA replication; cell division; growth;

KW proliferation; Candida albicans; fungicide; antifungal.

XX

OS Candida albicans.

XX

PN WO200253728-A2.

XX

PD 11-JUL-2002.

XX

PP 26-DEC-2001; 2001WO-US049486.

XX

PR 29-DEC-2000; 2000US-0259128P.

XX

PR 20-FEB-2001; 2001US-00792024.

XX

PR 22-AUG-2001; 2001US-0314050P.

XX

(ELT-) ELTRA PHARM INC.

XX

PI Roemer T., Jiang B., Boone C., Bussey H., Ohlseni KL;

XX

DR WPI: 2002-565694/60.

XX

N-PSDB: AB32388.

XX

Constructing strains for identifying gene products as effective targets for therapeutic intervention, by inactivating in the strain one allele of a gene and placing other allele of the gene under conditional expression.

Claim 44: SEQ ID NO 7675; 167pp + Sequence Listing: English.

XX

The invention relates to constructing (M1) a strain of diploid fungal cells in which both alleles of a gene are modified, comprising modifying one allele by insertion or replacement by cassette having an expressible selectable marker and modifying other allele by recombination, of a promoter replacement fragment with a heterologous promoter, so that expression of the second allele is regulated by the promoter. (M1) is useful for constructing a strain of diploid fungal cells in which both alleles of a gene are modified. The diploid fungal cells having both alleles modified are useful for identifying a gene that is essential to the survival or growth of a fungus, a gene that contributes to the virulence and/or pathogenicity of a fungus, a gene that contributes to the resistance of a diploid fungus to an antifungal agent, an antifungal agent that inhibits the growth of a diploid fungus and for identifying a therapeutic agent for treatment of a mammalian disease. (M1) is useful for identifying a compound which modulates the activity of a gene product, preferably enzymatic activity, carbon compound catabolism, biosynthesis, transcription, translational, signal transduction, DNA replication and cell division activity. The method is useful for identifying a compound having the ability to inhibit growth or proliferation of C. albicans cells and for

treating infection by *C. albicans*. The present sequence is that of an essential *Candida albicans* protein used in the method of the invention.

Note: The sequence data for this patent is not represented in the printed specification but is based on sequence information supplied to Derwent by the European Patent Office

XX
SQ Sequence 922 AA;

Query Match 99.9%; Score 4814; DB 5; Length 922;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 921; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 841 VPGEGGTNLGMFFIVGVLLVPLIPIIGSPFIAQWMLKLNNSIECQNNNGRPINHSSRSI 900
 Qy RSLGKXKHGKNSLISPPNKE 922
 Db 901 RSLGKXKHGKNSLISPPNKB 922

RESULT 3
 AAW07872 ID AAW07872 standard; protein: 859 AA.
 XX
 AC AAW07872;
 DT 23-FEB-1997 (first entry)
 DE Aluminum resistance gene ALR1.
 KW Aluminum resistance; ALR1; ALR2; ARH1; cation; tolerance;
 KW magnesium transport.
 XX
 OS Sacccharomyces cerevisiae.
 XX
 PN W0963459-A1.
 XX
 PD 07-NOV-1996.
 XX
 PF 01-MAY-1996; 96WO-NZ000035.
 XX
 PR 01-MAY-1995; 95NZ-0072319.
 XX
 PA (AUCK-) AUCKLAND UNISERVICES LTD.
 XX
 PI Gardner RC, Macdiarmid CW, Hay RJM;
 XX
 DR WPI: 1996-50161/50.
 XX
 PT ALR1-homologues, aluminium resistance genes ALR1 and ALR2 - used in the treatment of cation deficiency, cation toxicity, esp. heart disease.
 PS Claim 6; Fig 5; 62pp; English.
 XX
 CC ALR1 and ALR2 are isolated by selecting vectors which confer high aluminium tolerance, or that complement knockout mutations in ALR1, ALR2 and/or ARH1. The genes contain bacterial genes responsible for divalent ion uptake. The ALR1 and ALR2 are magnesium transport genes. The isolated transport genes can be used in the treatment of any plant, animal or microorganism disease which results from a cation deficiency, by producing an accumulation of cations in a plant deficient to those cations, or in plants consumed by animals deficient in those cations. The genes may also be used in the treatment of diseases resulting from cation toxicity, esp. manganese toxicity and esp. for the treatment of heart disease. Note: The ALR1 and ALR2 gene sequences are claimed and stated as given in Figure 6 and Figure 7 of the specification, however these figures are missing from the specification

SQ Sequence 859 AA;

Query Match 34.8%; Score 1679; DB 2; Length 859;
 Best Local Similarity 41.5%; Pred. No. 6, 2e-19;
 Matches 405; Conservative 137; Mismatches 258; Indels 176; Gaps 31;

Qy 1 MSDSESYTNTQSPRSPDEVDDHRNQTINDCAISDSELEKSLSESEVTKSKQ 60
 Db 1 MSDSESYTNTQSPRSPDEVDDHRNQTINDCAISDSELEKSLSESEVVKSKQ 60
 Qy 61 QHHQHITSDAKPLTRKGSSKKSNLTDKDRITNPMSLSSGGDPTTNSGHKNNYNTMS 120
 Db 61 QHHQHITSDAKPLTRKGSSKKSNLTDKDRITNPMSLSSGGDPTTNSGHKNNYNTMS 120
 Qy 121 SLRKOFYKLKDNTDNSTNNETHLAIPPIPIPITNNANKSERRKSQLENLPLIKKTPI 180
 Db 121 SLRKOFYKLKDNTDNSTNNETHLAIPPIPIPITNNANKSERRKSQLENLPLIKKTPI 180
 Qy 181 GRNNNFENDLVSPIAKTAKTNOSEDDTNTSTANNEKLGIGATTGCTGTATATA 240
 Db 181 GRNNNFENDLVSPIAKTAKTNOSEDDTNTSTANNEKLGIGATTGCTGTATATA 240
 Qy 241 AAGRGPSSSIDSEADSHASRSIDESEVDDCFPMVGDHIRVNGIDDETFBREBRA 300
 Db 241 AAGRGPSSSIDSEADSHASRSIDESEVDDCFPMVGDHIRVNGIDDETFBREBRA 300
 Qy 301 YLQKOMIAKNTLRIDEFONLSKNTTSGASRHYPFHHSNNNNKNNNGDGOGGSSMALEYTT 360
 Db 301 YLQKOMIAKNTNRSIDEFONLSKNTTSGASRHYPFHHSNNNNKNNNGDGOGGSSMALEYTT 360
 Qy 361 PKNLKLTKLSPREPHTNSSESEVYLTKHQPIPKYDQLSLSSTSSTSGSGSGYK 420
 Db 361 PKNLKLTKLSPREPHTNSSESEVYLTKHQPIPKYDQLSLSSTSSTSGSGSGYK 420
 Qy 421 PGAGATISDINGSPDRPSLSESETHADPISLYSPGSYVSDRLPRNGBETWMDLC 480
 Db 421 PGAGATISDINGSPDRPSLSESETHADPISLYSPGSYVSDRLPRNGBETWMDLC 480
 Qy 481 TGPDSSEMKAFLAKPGTHPLTAEDIHQETREKVELPKSKSYMPFCPHTFADEKESDYLEP 540
 Db 481 TGPDSSEMKAFLAKPGTHPLTAEDIHQETREKVELPKSKSYMPFCPHTFADEKESDYLEP 540
 Qy 541 INVYIVPHDGLTHPHSPISHPANVRRTYRDLVDYSDADMCLYALIDETDGFPAVYI 600
 Db 541 INVYIVPHDGLTHPHSPISHPANVRRTYRDLVDYSDADMCLYALIDETDGFPAVYI 600
 Qy 601 HQIYEADAIIDAIFDADTFDSMMLQIGESEKRYPATMLLISGADYTKMFARCOB 660
 Db 601 HQIYEADAIIDAIFDADTFDSMMLQIGESEKRYPATMLLISGADYTKMFARCOB 660
 Qy 661 ANSSGGYQRQYNIQQQQQAAPPNNPPIITSPNISTLNANSGTTGEGVGGTNGGP 720
 Db 661 ANSSGGYQRQYNIQQQQQAAPPNNPPIITSPNISTLNANSGTTGEGVGGTNGGP 720
 Qy 721 NPGGNNTNTNTGSPSPPOOQHITMKSPPIDARPRAFDIALYLGDIODHITMFQ 780
 Db 721 NPGGNNTNTNTGSPSPPOOQHITMKSPPIDARPRAFDIALYLGDIODHITMFQ 780
 Qy 781 NLLAYEKIFSRSNSYLAQLOVESFSNNNKITEMFSKILIGTMVPLNLVTLGFMNVR 840
 Db 781 NLLAYEKIFSRSNSYLAQLOVESFSNNNKITEMFSKILIGTMVPLNLVTLGFMNVR 840
 Qy 841 VPGEGGTNLGMFFIVGVLLVPLIPIIGSPFIAQWMLKLNNSIECQNNNGRPINHSSRSI 900
 Qy ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

| | | |
|----|--------------|---|
| PR | 31-AUG-2001: | 2001US-0316362P. |
| XX | PA | (ELITR) ELITRA PHARM INC. |
| XX | PT | New purified or isolated nucleic acids of essential genes of Aspergillus fumigatus, useful for treating or preventing infections by A. fumigatus, or for treating a non-infectious disease in a subject e.g. cancer. |
| XX | DR | WPI: 2003-093124/08. |
| XX | PI | Jiang B., Tishkoff D., Zamudio C., Broshkin AM., Hu W., Lemieux SM; |
| XX | PT | The invention relates to novel purified or isolated nucleic acids of essential genes of Aspergillus fumigatus. The isolated nucleic acids of the invention are used to treat or prevent infections by a pathogenic organism such as A. fumigatus, to treat a non-infectious disease in a subject (e.g. cancer), to prevent or contain contamination of an object by A. fumigatus, or to prevent or inhibit formation on a surface of a biolum comprising A. fumigatus. The polynucleotides are useful for expressing recombinant protein for characterisation, screening or therapeutic use, as markers for host tissues in which the pathogenic organisms invade or reside, for comparing with the DNA sequence of A. fumigatus to identify duplicated genes or paralogues having the same or similar biochemical activity and/or function, for comparing with DNA sequences of other related or distant pathogenic organisms to identify potential orthologous essential or virulence genes, for selecting and making oligomers for attachment to a nucleic acid array for examination of expression patterns, for raising anti-protein antibodies, as an antigen to raise anti-DNA antibodies or to elicit another immune response, and for identifying polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction. The polypeptides may be used to raise antibodies or to elicit immune response, as a reagent in assays designed to quantitatively determine levels of the protein in biological fluids, as a marker for host tissues in which pathogenic organism invade or reside, and to isolate correlative receptors or ligands in the case or virulence factors. This sequence represents a protein of one of the essential genes of Aspergillus fumigatus of the invention |
| XX | PS | Disclosure: Page: 175pp; English. |
| XX | CC | The invention relates to novel purified or isolated nucleic acids of essential genes of Aspergillus fumigatus. The isolated nucleic acids of the invention are used to treat or prevent infections by a pathogenic organism such as A. fumigatus, to treat a non-infectious disease in a subject (e.g. cancer), to prevent or contain contamination of an object by A. fumigatus, or to prevent or inhibit formation on a surface of a biolum comprising A. fumigatus. The polynucleotides are useful for expressing recombinant protein for characterisation, screening or therapeutic use, as markers for host tissues in which the pathogenic organisms invade or reside, for comparing with the DNA sequence of A. fumigatus to identify duplicated genes or paralogues having the same or similar biochemical activity and/or function, for comparing with DNA sequences of other related or distant pathogenic organisms to identify potential orthologous essential or virulence genes, for selecting and making oligomers for attachment to a nucleic acid array for examination of expression patterns, for raising anti-protein antibodies, as an antigen to raise anti-DNA antibodies or to elicit another immune response, and for identifying polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction. The polypeptides may be used to raise antibodies or to elicit immune response, as a reagent in assays designed to quantitatively determine levels of the protein in biological fluids, as a marker for host tissues in which pathogenic organism invade or reside, and to isolate correlative receptors or ligands in the case or virulence factors. This sequence represents a protein of one of the essential genes of Aspergillus fumigatus of the invention |
| XX | SQ | Sequence 663 AA; |
| XX | Query | Query Match 22.8%; Score 109; DB 6; Length 663; Best Local Similarity 32.7%; Pred. No. 9.7-e 251; Matches 292; Conservative 98; Mismatches 241; Indels 262; Gaps 22; |
| Db | Qy | 4 SESYSTNTNQIPRSDEVLDLQRHNRQTNDCAISDSEDELEKSELSEVSKBQQHQH 63 |
| Db | Qy | 7 SEKYMLCTRSTYVPE---LDDHRLQQL---DSPPRBATADI-SLSRQNTAQHAY 54 |
| Db | Qy | 64 HQEITSDNAKPLTRKSSSIIKKSNLTDKDRITNPMMSLGDDTINSCKRNNTMSSLR 123 |
| Db | Qy | 55 HQETPQ-----RPDLPSIQLAREGSLSRDFEQAILD 87 |
| Db | Qy | 124 KDFYLKD-----WTDNSNTNNH----THLAIPIPPIPITINMANKSRK---SQ 167 |
| Db | Qy | 88 DDRSGKDINTLGRFPSVDPNGVTRGRTWSRTHQL-----ANMSESSPSA 134 |
| Db | Qy | 168 LENSPPLIKKTTGRNNMSNPENDLVSPTRKAKTNDSEDIINTSTTAHKKLGIGATTLG 227 |
| Db | Qy | 135 RSSPP-----NSVFAADPERRRANTL-----SHAAFDLSEAQLQRTVSG 176 |
| Db | Qy | 228 VGTGT-TATATAAGAGRPSSIDEADSHASRSOETBD/CPPRVGHD/IRVNGIDF 286 |
| Db | Qy | 177 ---GTHPRPFSNASTARPQPDQIQLDN-----DESCVPPTEQGRIPVYD 222 |
| Db | Qy | 287 DEIDFIREERAYLQKQKQKLNKLNRIDEFNLSKNNTGASRSH-----THHSNNNNK 343 |
| Db | Qy | 223 EELFETVA-----LSRQNEFSTSR-----KQSLSQSSRGPRVFTDLPGLURKS 266 |

| | | | |
|---|--|--|---|
| Db | 305 | -----TVEKLQNENEPPTPQFPSSSESOSTHAAELCDLVPGD 342 | CC second protein, or its derivative, fragment, homologue or variant. The CC proteins are selected from given protein complexes, which are not defined in the specification. The variants are encoded by nucleic acids that CC hybridize to the nucleic acids encoding the proteins under low stringency CC conditions. The protein complexes are useful as targets for an active agent of a pharmaceutical. These protein complexes are particularly useful as drug targets for the treatment or preventing of a disease or disorder. The complexes and methods above are useful in diagnosing or screening for the presence of a disease or disorder or a predisposition for developing a disease or disorder in a subject. These are also useful in screening for a drug for treatment or prevention of a disease or disorder. The molecule that modulates the amount, activity or protein components of the complex is useful for the manufacture of a medicament for the treatment or prevention of a disease or disorder. This sequence corresponds to a protein of the invention. (Note: the sequence data for this patent did not form part of the printed specification but was obtained from the EPO in electronic format). |
| Qy | 464 | SVRDLFRNGEE--TWLDCCTPTDSEKMLAKAFGTHPLTAEDIRQETREKVELKSTY 521 | XX |
| Db | 343 | TFDLSQLQBGVWMLDVNLNEETBEAULSRSASIRLTDITLQEAEKVELFKQTY 402 | XX |
| Qy | 522 | FVCPTHTP-BADKESEDYLEPTINVVIVPHDGLTTHPSPISHPANTVRVRVQLRDYDV 580 | CC |
| Qy | 403 | FVCPTHTP-BADKESEDYLEPTINVVIVPHDGLTTHPSPISHPANTVRVRVQLRDYDV 462 | CC |
| Db | 581 | ADMIVCALVEDDETDGPAPVHGIEVAEATEDAYATPARDTDFSSMLQRIGESERRKWMUTM 640 | CC |
| Qy | 463 | SDWICYTAMIDIVDSFGPVIREBTEAIDLFLIAVRVDDFSFSPRIGLRKKYMSLM 522 | CC |
| Qy | 641 | RLLSGRADYKAFPARCOPBEANSSSGGTQRYONNQQQQQAPPNNPNTSPINSTLNL 700 | CC |
| Db | 523 | RLLSGRADYKAFPARCOPBEANSSSGGTQRYONNQQQQQAPPNNPNTSPINSTLNL 541 | CC |
| Qy | 701 | NSLGSTSTGGVGVGGINFGPNPFTGANTNTNTGSPSP0000QHGTNKSPPIDARP 760 | XX |
| Db | 542 | -----P 547 | Sequence 969 AA; |
| Qy | 761 | RADIALYLGIDODHITMFQNLLAYERKIFSRSHSYNLAQLOVESFNINNKITEMFSKITL 820 | Query Match |
| Db | 548 | RGDIGLYLSDIDQDHVUTMNSLAIFERMLSERSHSYNLAQLNVTNLULGHNVNVLKVSKVTL 607 | Best Local Similarity 24.28; Fred. No. 1.3e-547; Mismatches 271; Gaps 29; Matches 271; Conservative 143; Indels 460; |
| Qy | 821 | IGTMALVPLNLVTLGFMGNVRYVPGEGTNLGFEGVCFIIFTGSPFAQMW 873 | Qy 11 STTNOPIBRSDEVLDLHRNQLTNDCAISDSEDELXKLESEEVREVERSE----- 58 |
| Db | 608 | IATMLVPAHLIGLFGMANVRYVPGEGTNLGFEGVCFIIFTGSPFAQMW 660 | Db 44 SDSRSRPTOLIDNQLQHNGOTIDEDQI-DSWGMLHSDTSNDIKTSEDPSLKGAFIDHR 102 |
| Qy | 7 | ADK62466 standard; protein: 969 AA. | Qy 59 ----KQOORHQBITSD-NAREFLTRKSSSIIKKSNUYDKRITNPM-SLGGDDTINSCH 112 |
| AC | ADK62466; | XX | Db 103 PEMSQPREGPQSVSSTVQPOFINKFSPSPYKAGLRSQNSVLSUDSPSE--LESWL 160 |
| XX | 06-MAY-2004 (first entry) | XX | Qy 113 KNRNTNMSSLRKDYLKDN-TDNISTNRPHTHALAPIPIPPIPITNA----NRSRK 165 |
| DT | Disease treating protein complex-derived protein #359. | XX | Db 161 KER----KSVHKSTVDENSPDTRQSNIANNDVYD-VDMHVNNASTGYNDNSKRRK 215 |
| OS | protein complex; drug target; diagnosis. | XX | Qy 166 SOLENLPPLIKKKTIGRNNNNPNDLUSPMTMKINDSIDIINTSTTANHMKLGATT 225 |
| Unidentified. | | XX | Db 216 -----KKR-GSDSSN-KN--SKSTSDDNDEEDNTS----- 244 |
| PN | EP1338608-A2. | XX | Qy 226 LGVGTGTTATAAAGRFPSIDSEADSASSQTEEDVCPM--YGDITRYNG 283 |
| PA | (CBLL-) CELLZONE AG. | XX | Db 245 -----RFSSS-----LSNNNSSLDDCVLUDDESEPKAW 275 |
| PD | 27-AUG-2003. | XX | Qy 284 IDPFDIDEPFREEREAYLQKQIAKMKLRLIDFQNLSKNNTISGASHRHPYHHHSNNKK 343 |
| XX | | XX | Db 276 PDCTVLEPSSTEYER - LRS QAI-----ODARAIFIQYDDEZ 312 |
| EP | 20-DEC-2002; 2002EP-00102902. | XX | Qy 344 NNGGGCGGS SMAALKTPKNNLKKLTSRPFETHENSSEPEIYKLTKQPPYKDQLS 403 |
| XX | 20-DEC-2001; 2001EP-00130253. | XX | Db 313 D-----GTSNNDGTLFSKPKTVNTIDYELGRNRYNETENLKNGRPKRIAWHJLQIPM 367 |
| PR | | XX | Qy 404 LTSSISSTSSGSSQVQFSGARISDINGSL-----2D- RPSLFHSES 446 |
| XX | | XX | Db 368 VL-----GSNSTDSSRTSOSGLQDNLLVGRNTQYPHIIISNNPHEFRPTYRVL 418 |
| FA | WPI: 2003-638460/61. | XX | Qy 447 EETHADPPLPSGQSYDLP-----ADKES ----- 469 |
| (CBLL-) | DR-PSDB; AD62467. | XX | Db 419 DSTVHSPTISCLQPGOKFDPVASTYSDNSAGHIKTHPNSPTGIKATAVYVSOQGLT 478 |
| DR | | XX | Qy 470 -----RNGEET--WLDCTCPDSEKMLAKAFGTHPLTAEDIRNQETREKVB 515 |
| N-PSDB; | | XX | Db 479 AKNPSTLSSMSVANIEDVPPFWLDSNPIEEEEKKILSKAFGTHPLTBDIFGEYREVKE 538 |
| New proteins and protein complexes from eukaryotes, useful as targets in | | XX | Qy 516 LPKSYYPVCPHTFB-----ADKES ----- 534 |
| Drug screening, or in diagnosing or screening for the presence of a | | XX | Db 539 LFRDYLCIPSPDIVAEKHMRRKEQESATLDEIESRSRKSDQAYGATMSNESNANN 598 |
| disease or disorder, or a predisposition for developing a disease or disorder in a subject. | | XX | Qy 535 ----- 534 |
| Disclosure: SEQ ID NO 717; 13pp; English. | | XX | |
| The invention relates to novel protein complexes comprising a first and a | | CC | |

| | | |
|-----|--|-------|
| 599 | NSTNSARSKEWPLSIURARRSSANNTTTSSTYKRVSKRKBNEERFKRSGRDH | 658 |
| 535 | --EDYLEPINYIVYVPHDOLTEHFSPISSHANYVRVQLRDYDVDSADMUCYALIDE | 591 |
| 659 | KPRCGBEPIEPLYIVYVRCVLTHFAPTPHINVRARLKDLYANTISDIAVALIDD | 718 |
| 592 | ITDGCPAVHGEYEADAEDAVPTARDDESS----- | 624 |
| 719 | ITDAFAPMIELDEVYEATEDAALKHQSDDSDSSDGSASDDEAAPPFDVYSRK | 778 |
| 625 | -----MLQRIGESRRVYMTLMRLLSKADVI | 650 |
| 779 | TSYSSAKSEVSRSNSSTSASPANALIGNKRGDMURRIEGCRKVMSTILRLLSKADVI | 838 |
| 651 | KSPAKRCQEABANSYYQRTNLOOOQAOAPPNPNTISPINTNLNSLGSTGQQ | 710 |
| 839 | KGPAAK-----YEQW----- | 849 |
| 711 | VGGINFGENPQTNTNTGSPSPQQOQGINTNKSFPTDARPADIALYGD | 770 |
| 850 | ----- | ----- |
| 771 | IOPDHITMFONLLAYEKPSRSHSNHYLAQLOVESFNSNNKITEMSKITLICMVLPLNL | 830 |
| 864 | IOPDHITMFYSSLNHTKLRSHSNLYAQNIDMVKNDMDVQGFTILOTVPLPRN | 923 |
| 831 | VTCGLPFRGTRVPRPEBGTNLGMFPTGVLFITLGSFPAQ | 871 |
| 924 | ITSLDWRGTRVPIVQYQDSTLWTGIVLPRCMCLACSAVYHT | 964 |

KW thrombolytic; tumour inhibition; bodily characteristic; fertility;
 KW behaviour; cancer; proliferative disorder; neurological disorder;
 KW cardiovascular disease; immune system disorder; organ transplantation;
 KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
 KW hypothyroidism; cholesterol ester storage disease; infection; vulnerable;
 KW vasoconstrictor; antidiabetic; cytostatic; nootropic;
 KW neuroprotective; antiatherosclerotic; antiangiulant; thrombolytic;
 KW cardiotonic; hypotensive; antithyroid; antiinflammatory; immunomodulator;
 KW dermatological; analgesic; virucide; antibacterial; fungicide.
 XX OS Homo sapiens.
 XX EP WO200190366-A2.
 XX PD 28-NOV-2001.
 XX BR 24-MAY-2001; 2001WO-US017076.
 XX PF 24-MAY-2000; 2000US0-0206690P.
 XX ID (CURA-) CURAGEN CORP.
 XX PA (CURA-) CURAGEN CORP.
 XX PI Leach MD, Shinkets RA.
 XX DR WPI; 2002-106200/14.
 XX DR N-PSDB; ABN76186.
 PR Novel human polypeptides and polynucleotides useful for diagnosing,
 PR preventing and treating cardiovascular diseases, neurodegenerative,
 PR hyperproliferative disorders and disorders related to organ
 PR transplantation.
 XX ES Claim 10; Page 830; 2508pp; English.
 CC Sequences ABP31028-ABP35561 represent 4534 novel human proteins
 CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
 CC ABN7587 represent cDNAs encoding them. The invention also encompasses
 CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
 CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
 CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
 CC polynucleotides, the recombinant production of ORFX proteins, antibodies
 CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
 CC polypeptides, methods of screening for modulators of ORFX expression or
 CC activity, and methods of screening individuals for a predisposition to an
 CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
 CC range of biological activities, such as cytokine, cell proliferation,
 CC cell differentiation, immune modulation, haemopoiesis regulation,
 CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
 CC chemokinetic activity, haemostatic activity, thrombolytic activity,
 CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
 CC and antinefetive activity, and may also be involved in the determination
 CC of bodily characteristics, fertility and behaviour. ORFX proteins,
 CC nucleic acids and antibodies may also be used as a
 CC source of primers and probes, in the detection of ORFX genomic sequences
 CC or transcripts, in the identification and cloning of homologous
 CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
 CC nucleic acids may additionally be used to produce transgenic animals
 CC which may be useful for studying the function and/or activity of ORFX
 CC protein, and in drug screening. The ORFX proteins may also be used as
 CC immunogens to generate specific antibodies, which are useful in the
 CC diagnosis, treatment and monitoring of ORFX-associated diseases
 XX SQ Sequence 154 AA;

| | | | | |
|--|---|-----------------|------------|-------------|
| Query Match | 7.2% | Score 345.5; | DB 5; | Length 154; |
| Best Local Similarity | 44.6% | Pred. No. 5e18; | | |
| Matches 79; | Conservative 20; | Mismatches 39; | Indels 39; | Gaps 6; |
| Qy | 351 GSSMAALKYTPKNIKTKTLSPFETHNSSSSEIYELTKIQOPPKYDQLSLTSSTSS 410 | | | |
| Db | 1 GSS-ALKYTP--ISR--TELEKNCSTVNTYSENNS-VREDDKPDL---- 42 | | | |
| Qy | 411 TSGSGCGVFGGARISDGING-----SLPDRFLPFSSEBETHADP 454 | | | |
| Db | 43 ----HPDPVPGRAKIGEGRGNDSYSRAYTQLQNTEQIPSRSPFSEDETASD 97 | | | |
| Qy | 455 IPSLIVSPGQSYSTRDLFRNGETWLDCTCPDSDERKMLAKAFGTHPLTADIMQTR 511 | | | |
| Db | 98 IPSLISBQQTYYELFKGQDPWWLDSCSPTDDEMRCIATKFGTHPLTADIRMQTR 154 | | | |
| <hr/> | | | | |
| RESULT 10 | | | | |
| ID ABP06753 | standard; protein; 107 AA. | | | |
| XX | | | | |
| XX | | | | |
| AC ABP06753; | | | | |
| XX | | | | |
| XX | | | | |
| DT 24-JUN-2002 (first entry) | | | | |
| XX | | | | |
| DE Human ORFX protein sequence SEQ ID NO:13488. | | | | |
| XX | | | | |
| KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis; KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage; KW degenerative disorder; osteoarthritis; neurodegenerative disorder; KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus; KW hypertension; hypothyroidism; cholesterol ester storage disease; KW immune deficiency; immune disorder; infectious disease; KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis; KW myasthenia gravis. | | | | |
| XX | | | | |
| OS Homo sapiens. | | | | |
| XX | | | | |
| PN WO200192521-A2. | | | | |
| XX | | | | |
| CC 06-DEC-2001. | | | | |
| CC 29-MAY-2001; 2001WO-US010836. | | | | |
| CC 30-MAY-2000; 2000US-0206112P. | | | | |
| CC 29-AUG-2000; 2000US-0228716P. | | | | |
| XX | | | | |
| PA (CUBA-) CURAGEN CORP. | | | | |
| XX | | | | |
| PI Shinkets RA, Leach MD; | | | | |
| XX | | | | |
| DR WPI; 2002-106308/14. | | | | |
| DR N-PSDB; ABN22505. | | | | |
| XX | | | | |
| CC The present invention describes substantially purified human proteins CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1 CC in the specification). ABN15762 to ABN2732 encode the human ORFX CC proteins given in ABP0001 to ABP11500. ORFX proteins are useful for CC treating or preventing a pathology associated with an ORFX-associated CC disorder in humans, and in the manufacture of a medicament for treating a CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide CC sequences can be used in gene therapy. ORFX sequences can be used in the CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver, | | | | |
| PS Disclosure; SEQ ID NO 13488; 1037pp; English. | | | | |
| XX | | | | |
| PT Novel human polypeptides and polynucleotides useful for diagnosing, PT preventing and treating cardiovascular disease, neurodegenerative, PT hyperproliferative disorders and autoimmune disorders. | | | | |
| XX | | | | |

psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage, osteoarthritis, neurodegenerative disorder, disorders related to organ transplantation, cardiovascular diseases, diabetes mellitus, systemic lupus erythematosus, hypertension, hypothyroidism, cholesterol ester storage disease, various immune deficiencies and disorders, infectious diseases, autoimmune disorders such as multiple sclerosis, rheumatoid arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. ORPX proteins are also useful for treating burns, incisions, ulcers, for treating osteoporosis, bone degenerative disorders, or periodontal disease, and for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues and conditions resulting from systemic cytokine damage. N.B. The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form directly from WIPO at <http://wipo.int/pct/publication/pct-sequences>

XX Sequence 107 AA:

| | |
|-----------------------|---|
| Query Match | Score 319.5; DB 5; Length 107; |
| Best Local Similarity | 6.6%; Pred. No. 2.8e-16; |
| Matches | 56; Conservative 21; Mismatches 11; Indels 1; Gaps 1; |

Qy 790 SRSHTNTLAQLOVESFNSNNKTEPFSKTTLIGTMLVPLNIVTGFLGMYRVPGEGGTNL 849

Dy 1 SRSHTNTLAQLOVESFNSNNKTEPFSKTTLIGTMLVPLNIVTGFLGMYRVPGEGGTNL 59

Qy 850 GMPFGTVGVLIPIITGSPIAQWNIKKLN 878

Dy 60 AWWFGILGVLLLAIVLGFLASWIKRID 88

RESULT 11

ID ARN92441 standard; protein; 326 AA.

XX AC ARN92441;

XX DT 02-JUN-2005 (first entry)

DB M. xanthus protein sequence, seq id 11640.

XX FW Transgenic plant; DNA replication; gene regulation; gene expression.

XX OS *Myxococcus xanthus*.

PN US6833447-B1.

XX PD 21-DEC-2004.

XX PP 10-JUL-2001; 2001US-00902540.

XX PR 10-JUL-2000; 2000US-0217881P.

XX PA (MONS) MONSANTO TECHNOLOGY LLC.

XX PI Goldman BS, Hinkle GJ, Slater SC, Wiegand RC;

XX DR WPI; 2005-028716/03.

XX PFT New substantially purified *Myxococcus xanthus* nucleic acid molecule encoding a nitrite reductase, useful for determining gene expression, identifying mutations in a gene of interest, and for constructing mutations in a gene of interest.

XX PS Example 2; SEQ ID NO 11640; 25pp; English.

XX CC The invention relates to a substantially purified nucleic acid molecule encoding a nitrite reductase of SEQ ID NO. 11640. Further disclosed is a recombinant DNA construct for expression of a nitrite reductase gene in a plant cell, and a plant cell comprising the recombinant DNA construct.

CC The nucleic acid is useful for determining gene expression, identifying mutations in a gene of interest, and for constructing mutations in a gene of interest. Sequences given in records for SEQ IDs 962-16825 represent a group of 7114 *Myxococcus xanthus* proteins and peptides. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO

XX Sequence 326 AA:

| | |
|-----------------------|---|
| Query Match | Score 251; DB 9; Length 326; |
| Best Local Similarity | 20.8%; Pred. No. 2.5e-10; |
| Matches | 90; Conservative 63; Mismatches 136; Indels 144; Gaps 12; |

Qy 458 LVSPGQSVDRDLFRNGEETW-----WLDCTCPDSEKMKLAKAAGFTHPLTAEDIRHQETR 511

Dy 18 LVKDQQ---LITGEELBEGIEKGKWKDVHPTBEVRLAAGHKLAVEDCLLDDOR 73

Qy 512 EKVLFPSYXNCFCHTFPAKDESDYLEPINVYVVFHDOLTPSPISHPANYPRRYR 571

Dy 74 PKLEEFPHNQFQVVLQGFETVTKDICELTLEHHFLAKDWLPHGHTVRRVK 133

Qy 572 QRDYDVS-----DUCLVALIDEITDOPAPVIGCIETEADAKIEDAYFTADDFTSSEN 625

Dy 134 E-----DPAQTLIRGTDQFDVFTMLADALVDAQFLDTSSEBDLAAIF--DKEPKSH 185

Qy 626 LORIGEBSRKWTMFLRSLGADYVKAHPKRCOBANEASSGGYQRQYQNLQOOQQAPPBP 685

Dy 186 LQITAFMRMIVTFRVLSQDYGULSER-----216

Qy 686 PNPIITSPINSTLNLSLTGTCGGVGQGGINFGPNETGANTNTNTGSPSPQQQQQ 745

Dy 217 -----GI-----SHVHEKTT-----226

Db 746 HGITMKNSPPIDPAPRADIALYLGDIODHITMFQNLLAYEKIFSRSHSNYLAQLQVESF 805

Qy 227 -----LYFPRDVTDHLYRVEQDARDIIGVIVGIVQYLSMVA----262

Dy 806 NSNNKITEMFSKTTLIGTMLVPLNIVTGFLGMYRVPGEGGTNLGMFP---GIVGVLFPI 862

Qy 263 ---NFTNDISQLTIPATIFPLSCIVGFFQNP---DRLYTNAYVLLWASTLGPVGM 316

Dy 863 IIGSPIPAQWNLK 875

Db 317 V-----FWFK 321

RESULT 12

ID ABB60371 standard; protein; 1198 AA.

XX AC ABB60371;

XX DT 26-MAR-2002 (first entry)

DB Drosophila melanogaster polypeptide SEQ ID NO 7905.

XX KW Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.

XX OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX DT 27-SEP-2001.

XX PP 23-MAR-2001; 2001WO-US009231.

CC XX

CC PR 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

CC XX

DB Lactococcus lactis protein yqfc. 616 TSGAPL-----SSNQTSSEASSNSMSIN-SPSLSLSLT-SNSBESATNQSNSSEAT 664

KW Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese. QY YELKTKQQPPYKVDQSLSSSSTSGSGGARISDGINGGSAF--DRFLSLTH 443

XX OS Lactococcus lactis; IL1403. DB 665 -----KVDNNTSTHSNILANGSDNDSPDSDDSDPDSNLSSSPNLETTQNTS 712

XX PN PR2807446-A1. QY 444 SESEETIHPDIPSIVPKVSSVQVQ-----CDPSEMMALAKAFTHULTR 503

XX PD 12-OCT-2001. DB 713 SKPSEVNNSPERPKVSSVQV-----CDPSEMMALAKAFTHULTR 503

XX PP 11-APR-2000; 2000FR-00004630. QY 504 DIRMOETRKVEFLKSYTFVCPHTEAKDESBDYL---EPINVYVVFHDGILTFHSP 559

XX PR 11-APR-2000; 2000FR-00004630. DB 759 SSSQO-----KESQNLNTTGIN-----NPITNNNS 787

XX PA (INRG) INRA INST NAT RECH AGRONOMIQUE. QY 560 TSHPANVRRVRQLRDYD--VSADWLCTALDEI-TDGFPAVIGIEYADEDAVFP 615

XX PI Bolotine A, Sorokine A, Renault P, Ehrlich SD; DB 788 SENSA----ASILTSVSNNSSESETCLYTNEAORDNGSISLSP-SENNSENNNTS 841

XX DR WPI; 2002-03418/06. QY 616 TARDDTDFSSMLQRIGESRRKUMTMRLLSGRAVDTKMPAKRCOBANESSGTYQROYNLQ 675

XX PT New nucleotide sequence useful in the identification or Lactococcus DB 842 SIQS---QALESSKSSTNKRSULLSINS-----TSHPONEDQNSSEDEVSNNN 890

PT lactis and related species. QY 676 QQQQQAPPPPNNPITSPINSTLNLSIGTSTGGYVGCGTNGPNTGNNTNTNTG 735

XX PS Claim 6: SEQ ID NO 1665; 2504pp; French. DB 891 VIES-----ILCOLNSLISKNTKTRANSLTSOKSVITYLPSKVKTNKEWNTNTS 939

CC The present invention is related to a Lactococcus lactis nucleotide sequence (AA90521) and related proteins (AB33300; AB05621). The nucleic acid sequence is useful in the detection and/or amplification of nucleic acid sequences, particularly to identify Lactococcus lactis or related species. The proteins of the invention are useful for the biosynthesis or biodegradation of a composition of interest. The invention helps research in lactic bacteria, particularly useful in the production of yogurt and cheese. Note: The sequence data for this patent is based on equivalent patent WO200177334 (published 18-OCT-2001) which is available in electronic format directly from WIPO at [ftp://wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences). (Updated on 29-AUG-2003 to standardise OS field)

XX SQ Sequence 1022 Aa: Search completed: June 13, 2006, 14:08:12

Query Match 3.9%; Score 186.5; DB 5; Length 1072;
Best Local Similarity 20.8%; Pred. No 0.00012;
Matches 193; Conservative 136; Mismatches 363; Indels 235; Gaps 45;

QY 1 MSDSESYYQ-NSTTNOPIPRSEDEVLDHHRNQIN-DCAISSDEDELKSELESEVYKSE 58

DB 305 VASSESTDANSALYPTSEASISYDNTANSISLSDSISSQTB-----NSQ 352

QY 59 KQQQHHEQETSDNAKPLTRKSSGSSKKSNLTDKDRITNPMMSLGGDDTINSGHKNRTN 118

DB 353 SGASSTABISTOSSENSLNSQNSNSN-SED--SNQSLG--SSMSNEESEN 406

QY 119 MSSLRDFKQKNTDINSHHHLAIPPIPPIPITRANSRKSQLENPPFLKK 178

DB 407 SENI-----NETNNNSP1-----TNIILPPSPNTENS-----VSDQ 437

QY 179 TIGRNNSNNFENDLVSPATKTKTIDSDIINTS----TTRANMKGIG-----222

DB 438 TSSEASTNSNSNSISSPNISSTSDDSATNSSPDNVAYAANSLASTNNSSSVLSST 497

QY 223 --ATTGIV----GTTTATATAAGRPSPRSSIDEADSHARSQTEBEDCPCMG 276

DB 493 SPADLGINQSGSDNLTUDSSEIESTSGAFLSNCTSSEATNSNSISLSPNISSTSVL 557

QY 277 DHRYNGIDPDSIDEFIR----BERBEAYLQKQIAKNTLRIDP--QNLSKNN---T 325

DB 558 ES-TMSSMENSAVANNNSLASTNNSSSVLSSTPSTAQN-GSDNLTKQSELS 615

QY 326 TSGASRHPYHHSNNNNKGDGCGSSMAALKYTPNKLKPLSRPEFTHENSSSSERI 385

| | | |
|----|--|---|
| DS | Debaryomyces hansenii (Yeast) (<i>Torulaspora hansenii</i>). Eukarya; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; | 453 VRGCGSVSDLFRNGCITWLDCCPCTDAEMKVMARAFHPLTAEDIRMOETRKEVLFPR 518 |
| IC | Saccharomycetales; Saccharomycetaceae; Debaryomyces. | |
| NX | NCBI_TaxID=4958; | |
| NA | [1] | |
| NP | NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA]. | |
| SP | STRAIN=ATCC 36239 / CDS 767; | |
| LP | PubMed:15220592; DOI:10.1038/nature02579; | |
| LA | Dujon B., Fischer G., Durrens P., Casaregola S., | |
| LA | Lafontaine I., de Montigny J., March C., Neuveglise C., Talla E., | |
| LA | Goffard N., Prangeul L., Aigle M., Anthouard V., Babour A., Barbe V., | |
| LA | Bourne S., Boyer J., Becker J.-M., Beyne E., Blaykasten C., | |
| LA | Boistram A., Blanchard S., Contantinoff P., de Daruvar A., | |
| LA | Despons L., Fabre B., Fairhead C., Ferry J.-D., Grappi A., | |
| LA | Hantraye P., Hennequin C., Jauriaux N., Joyet P., Kachouri R., | |
| LA | Karrest A., Koszul R., Lemire M., Lesur I., Ma L., Muller H., | |
| LA | Nicraud J.-M., Nikolski M., Ozier-Kalogeropoulos O., | |
| LA | Pellanz S., Potier S., Richard C.-P., Straub M.-L., Souleau A., | |
| LA | Senewi-Meyer M., Wesołowska-Louvel E., Westhof E., Wirth B., | |
| LA | Zeniou-Meyer M., Zivanovic Y., Bolotin-Fukuhara M., Thierry A., | |
| LA | Bouchier P., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J., | |
| LA | Wincker P., Souciet J.-L., | |
| LA | Genome evolution in yeasts."; | |
| LA | Nature 430:35-44(2004). | |
| SC | Copyright by the UniProt Consortium. See http://www.uniprot.org/terms | |
| DC | Distributed under the Creative Commons Attribution-NoDerivs License | |
| IC | ENBL: CH382138: CAG8451.1: - Genomic DNA. | |
| DR | GO: GO:0016020: C:membrane; IEA. | |
| DR | GO: GO:0046873: P:ion transporter activity; IEA. | |
| DR | GO: GO:0030001: P:metal ion transport; IEA. | |
| DR | InterPro: IPR002532: Mg2+-transptcorA. | |
| DR | PFam: PF01544: CorA_1. | |
| W | Complete proteome. | |
| W | SEQUENCE 936 AA: 104547 MW: 8F050BA9A42104 CRC64; | |
| W | Query Match 42.7%; Score 2056; DB 2; Length 936; | |
| W | Best Local Similarity 46.5%; Pred. No. 2.1e-108; Gaps 31; | |
| W | Matches 45; Conservative 108; Mismatches 223; Indels 226; | |
| W | 1 MDSDESYYCONSTINOPIRE----SDEVLEHDHQNT----DCAISDSEBE 45 | |
| W | 1 MSDNTESPRVSVPSPRETLIRAAANTGEVLNDHRNQLBSLPIRGTCGGMCOSSAVSDDDDDY 60 | |
| W | 46 LASELESEEVSKERKQOHQB E-----TSNDAKPLTRKSISIKRSNLTKDRIT 96 | |
| W | 61 QSSRSEDAQYQQQQQQPQPFPHQLOGSRSDVEQFPEAQQQCSQYRS----- 111 | |
| W | 97 NPMFLSGGDDTINSGHKNVNMMSSLRKDPYKONTDON-----STNNPHHLAIPPIP 150 | |
| W | 112 -----SKNAANLRLDFTLKNDDNERYTPSKNQPLQSSS-PVG 154 | |
| W | 151 IPTIPTTNNANK-----RRIKSQLENLP-----LHK----RKT 179 | |
| W | 155 RSGGCTGTAQPKQRSGRPQSOSPSGQAKSDGDEVRRHQAStAAAPPFGILMKUSNPPSKK 214 | |
| W | 180 IGRNNS---ANNENTDVLUSPMTKAKNSDSETITNTTANMKLGATLGVGRTATA 236 | |
| W | 215 SKRSNMINDTLEADLAAPMTSKINNS---MTSEANONNNK----- 252 | |
| W | 237 TATAAAGGRPSSSDPSADSHASHSSOETEDWCFCPVNGDHTIRYNGIDPEDEFIRE 296 | |
| W | 253 ---RERRPSITSIDSAHSASSQETEEDCPFLHVKKGIDPDEIEFIRDE 307 | |
| W | 297 RE-EAYL---QKOMIA-KNLRIDDEFQNLSKNNTGASRHPYHHNNNNKKNGDGEGS 352 | |
| W | 308 KENENHLKEEQMIAERTAMD-----T3HGSAARDV-----GP 341 | |
| W | 353 SMAALKTPKNIKLTKLTSRPEETHEN-----LKTQKOPPKY 398 | |
| W | 342 SKNALKTPKNIKLGA-SPGKGRNNTGAGDMSSGGINISERAQSDSSTVDDSSARY 400 | |
| W | 399 DPOLSLSTESTSSTSGSGCQYKPGCARIISDGINGSLPDRSLPHSSESETHADPSL 458 | |
| W | 401 NEKI_LGDDDISTE ---NYKFGTGTTRD---MSALBERTSPCSSESETHADPSL 452 | |
| W | 459 VSPGAGSVRDLFNGFBTWLDCCPTEPDSEMKAAGIHLPLTEDIRMOETRKEVLFK 518 | |
| CC | Copyrigted by the UniProt Consortium, see http://www.uniprot.org/terms | |
| CC | Distributed under the Creative Commons Attribution-NoDerivs License | |
| CC | EMBL: CR38051; CAG8451.1: - Genomic_DNA. | |
| DR | GO: GO:0016020; C:membrane; IEA. | |
| DR | GO: GO:0046873; P:metal ion transporter activity; IEA. | |
| DR | GO: GO:0030001; P:metal ion transport; IEA. | |
| DR | InterPro: IPR01544; CorA_1; Pfam: PF01544. | |
| DR | Nature 430:35-44(2004). | |
| CC | Copyrigted by the UniProt Consortium, see http://www.uniprot.org/terms | |
| CC | Distributed under the Creative Commons Attribution-NoDerivs License | |
| CC | EMBL: CR38051; CAG8451.1: - Genomic_DNA. | |
| DR | GO: GO:0016020; C:membrane; IEA. | |
| DR | GO: GO:0046873; P:metal ion transporter activity; IEA. | |
| DR | GO: GO:0030001; P:metal ion transport; IEA. | |
| DR | InterPro: IPR01544; CorA_1; Pfam: PF01544. | |
| DR | Nature 430:35-44(2004). | |

| | | | | |
|---|--------------------|-----------|-------------------------|--|
| RW | Complete proteome. | 89694 MW; | E741B17C22E235A4 CRC64; | |
| 50 | SEQUENCE 803 AA; | | | |
| Query Match 34.9% ; Score 1683.5 ; DB 2 ; Length 803 ; Best Local Similarity 43.0% ; Pred. No. 2.9-87 ; Matches 413 ; Conservative 92 ; Mismatches 215 ; Indels 241 ; Gaps 31 ; Matches 413 ; | | | | |
| Qy 38 SDESE---LELSELESEVSESERQQ----OHHQIQTSDNAKPLTRKSSSIKKKN 88 Db 3 SSESSEPHSPOLHQSVAHSKVNKTDEHNLVYDHRHIS----VPPRLGDSFGSSN 56 | | | | |
| Qy 89 LTKK----DRITNP---MSLGGDDTINSGHKRNTNMSSLRKFY-LKDNTDNSTNN 139 Db 57 LNARSHMMLDRAANEQPGTMNIEGADS----DCPFLGKDHKLRSALQNSMAR 105 | | | | |
| Qy 140 MTHLAPL--PIPIPPLIT-NANKSERKSQLENPLPKTIVGRNANSANFENDLVP 195 Db 106 PSRLAITYNHQEVQPATITDNEELKNSVAAIASPV----TKTRGSNTSFLTDKTP 160 | | | | |
| Qy 196 MTKMKTNDSEDDTNTSTTAHKRKGIGATTLGVTGTGTTATAAGGRPSRSSID--- Db 161 L-KRTMSRNINSAVSSNRLN-----AVNSISIHKSYDDN 199 | | | | |
| Qy 253 -----SEADSHASRSSCSSEETEDVCFPAVDG-HIRYNGIDFDEDFIREKEEAY 301 Db 200 LKKKTRTASSPSDNSQAERSCSEETEDVCFPMPPOLHRSRNGIDPELEPAAEVNE--- 256 | | | | |
| Qy 302 LOKOMIAKLNILRDFPONLSKNTTSGASRHPYHHSNNNNKNGGDGGSSK----- Db 257 -QRKLLC-----EISLAAKSG-----SNG-QAOSSSSSANSFRS 289 | | | | |
| Qy 355 -ALKYKPNLILKYL--RFPFTHESSSELYE-----LTKKQQPYKDDQ 401 Db 290 EAALKYKTPH--LRKOLSPKSESNNSNSNSPDPNOGVSFGNNTEGDED--YEAR 345 | | | | |
| Qy 402 LSITSSSTSSTSGSGCOVKEGGARISDGINGSPDPRSFPHSESESETHADPILVSP 461 Db 346 PQAYTOPENPQYGDPSK-----ADPRSFPSSEMMESTHATDLNVTLP 391 | | | | |
| Qy 462 GQVRDLPFNGBETWLDCTPDTSENKMLAKAFGIPHLTAEDIQFQETREKVLFKSTY 521 Db 392 GQSFPLFPGPPTWMLCSTPDTDENCRKAGPISPLAEDIQFQETREKVLFKSTY 451 | | | | |
| Qy 522 EWCETPDKSESDYELPINVYIVWHDGLATLPHESPISHRANVRVRQRLDVTDA 581 Db 452 FWCFTPDKSESDYELPINVYIVWHDGLATLPHESPISHRANVRVRQRLDVTDA 511 | | | | |
| Qy 582 DMECLALIDTGPATHGIGYEADAEVATARDTDFSMLORGSRSEKWTLMR 641 Db 512 DMCYALIDDTDSAPVTSRSELEASIDESVTPMDLFAMMLORIGSRSEKWTLMR 571 | | | | |
| Qy 642 LLSGRADYKTFPAKCOPEANSSGSSYQTYQYNLQQQQQPPPPPLITSPINTNLN 701 Db 572 LLSGRADYKTFPAKCOPEANSSGSSYQTYQYNLQQQQQPPPPPLITSPINTNLN 604 | | | | |
| Qy 702 SIGTSTGGVGQGINFGPNGTNTNTGSPSPPOQQQHQGITMKSPPIDPDRP 761 Db 605 -----ANLTSSSPNVKXMQKLQMN-----TQR 628 | | | | |
| Qy 762 ADIALYQIQHIIHTMTQNLAEKTFSRGHSA LAQLOVEFSNPKKTFSKVTLI 821 Db 629 GDIALYLGPIQHLLTUMYQNIATEKFSRSHTNLAQLOVEFSNPKKTFSKVTLI 688 | | | | |
| Qy 822 GTMLYPLNUTVGFEMMYTRP EGCTNLGMWFGFQVWLPIIIGSTPFAOMWLKK---- Db 689 GTVLPFLNUTVGFEMMYTRP EGCTNLGMWFGFQVWLPIIIGSTPFAOMWLKK----TQRL 747 | | | | |
| Qy 877 -LNNSTB-----SONNNGNPFIHNHSSRSIRSGLURKKG--NRSITSPPRK 920 Db 748 TUNEAESGAKSILSSFIPRQPNENDEKY-----SRGRSKYGGPSNRSLASLPSK 798 | | | | |
| Qy 921 Y 921 Db 799 Y 799 | | | | |
| RESULT 5 ARL1_YEAST ID ARL1_YEAST | STANDARD | PRT; | 859 AA. | |

| | | | |
|----|--|---|---|
| DR | Ensembl; YOL130W; <i>Saccharomyces cerevisiae</i> . | Db | 721 YLAQLQVESFSNNNKITEMSKITLGMVTPNVLVTRGKVNPKGE-NSSIAWWFGI 779 |
| DR | GenomeReviews; Y13140_GR; YOL130W. | Qy | 836 VGVLLIFTINGFPIAQWKLKNSI---HSERRSTRSGLRK 907 |
| DR | SGD; S00005490; ALR2. | Db | 780 LGVLLAVLGWLPLASWIKRIDEPATNEAASSGAKSVISSFLPERKRNFRDKSNTV 839 |
| DR | BioCyc; SCER-S28-01; SCER-S28-01-005413-MONOMER; -. | Qy | 908 HGG--NKSISFPK 921 |
| DR | LinkHub; Q08269; -. | Db | 840 RAGPSNKNSVSLPSRY 855 |
| DR | GO: 0005886; C: plasma membrane; IMP. | | |
| DR | GO: 0015082; F: di-, tri-valent inorganic cation transport; IDA. | | |
| DR | GO: 0015693; F: magnesium ion transport; IDA. | | |
| DR | InterPro; IPR012523; Mg+-transporter. | | |
| DR | Pfam; PF01544; CorA; 1. | | |
| RW | Complete proteome; Magnesium; Membrane; Transport. | | |
| FT | CHAIN 1 859 | | |
| FT | Magnesium transporter ALR2. | | |
| FT | PRO_0000201535. | | |
| FT | Potential. | | |
| FT | 744 764 | RESULT 6 | |
| FT | 773 793 | ALR2_YEAST | |
| FT | Potential. | STANDARD; | |
| FT | 13 13 | ID AC | |
| FT | N > Y (in Ref. 1). | P43531 | |
| QX | SEQUENCE 859 AA; 95869 MW; 6DA4ACA/0EFAC63 CRC64; | DT 01-NOV-1995, integrated into UniProtKB/Swiss-Prot. | |
| Qy | 1 MSDSESYQNSTNPIRSDEVLDIIRRQNITDCAISSEDBLEL--KSELESEWORKS 57 | DT 07-MAR-2006, entry version 40. | |
| Qy | 1 MSSSS---SSSENPLRS---NSLANTYMSMKTEDITYGIDYRQHIDSPVPRH 49 | DE Name=ALR2; OrdererdLocusNames=VF050C; | |
| Db | 58 ERQQQHQEQTDSNAKPLTRKSSSIKKK---NLTQDKRTNPMSLSSGDDTNSGH 112 | GN OS | |
| Qy | 50 QPPTLKNEKAKSTKPS-TPKQKSKATYNSHVDGTSVPSGRNDPDEQGQDDEVY-AHH 108 | NCBI_TAXID=4932; | |
| Db | 113 KOR---NYNMSSL---RKDYLKDNTDINSNNTH---HLA1PIPIPIPTI 156 | OX [1] | |
| Qy | 109 QIRASAILTSARPLAHNSHPHQRYESVNI-----HTTPKGTVKRDYTMSSSTA 161 | RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA]. | |
| Db | 157 TNAKSRSRKSQLENLPLPIKKTTGRNNNNFENDLVSPMTKHTKTDSEDTNT-----210 | RC STRAIN=S28C / AB972; | |
| Qy | 162 SSANKS-K1-SASSASAPTKV-----KSLVSPVLEPHHSKSDTHSKLAKPK 210 | RX MEDLINE=95400292; PubMed=1670463; | |
| Db | 211 ---SITANRMKIGCATPLTGTCGTTATA---TAAAGRGRPRSSTDASDHRSQ 264 | RA Magnesium transporter ALR2 (Aluminum resistance protein 2). | |
| Qy | 211 RTYSPTSAHSIN-PAVILUTSTSQSDADDYLTERKPVMTNTRAFDSDV-SQASRSQ 268 | RA Sacccharomyces cerevisiae (Baker's yeast). | |
| Db | 265 ETEEDVCPEMGD-HIRUNGLDEFIDEFIREEEAYLOKOMIAK-NILRIDEFONLSK 322 | RA Saccharomyces cerevisiae. | |
| Qy | 269 ETEEDVCPEMGD-HIRUNGLDEFIDEFIREEEAYLOKOMIAK-NILRIDEFONLSK 322 | RA Saccharomyces cerevisiae. | |
| Db | 323 -----NNNTGASRHPYTHHNNNNKNGDGGSSMAALKYTPKMKLKTLSREETH 376 | RA Dephoure N, O'Shea E.K., Weissman J.S.; RT Global analysis of protein expression in yeast.; | |
| Qy | 326 DIGFTSPTTSLOSS-----AALKYTPR--VSQGEKESTN 359 | RA Nature 425:737-741 (2003). | |
| Db | 377 ENSSSERETLYKIKTQOPPYKYYDQLSUTSSTSSTSGSSGKFGG-ARISDINGGSL 435 | CC --!- SUBCELLULAR LOCATION: Plasma membrane magnesium transporter. | |
| Qy | 386 ET-----KPSLQPNQEKV-----KPSLQPNQEKV-----KPSLQPNQEKV 414 | CC --!- POTENTIAL: Membrane: multi-pass membrane protein (Potential) | |
| Db | 436 PDRPLSPSESETHEADPISLVSQPSYRDLFNGEETWMDCPTDSEKTMKAF 495 | CC --!- MISCELLANOUS: Present with 2130 molecules/cell. | |
| Qy | 436 PDRPLSPSESETHEADPISLVSQPSYRDLFNGEETWMDCPTDSEKTMKAF 495 | CC --!- SIMILARITY: Belongs to the corA family. | |
| Db | 415 PNRFSPFCESEDETHASDIPSLSYSEQTYELFRGEPTWWLDCSCPDEMRCIAKAF 474 | CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms | |
| Qy | 496 GHPIIATDINMQETREKVELFKSYTYCPHTEAKEDSYLKEPLNMYVYFHDQIITP 555 | CC Distributed under the Creative Commons Attribution-NoDerivs License | |
| Db | 475 GHPIIATDINMQETREKVELFKSYTYCPHTEAKEDSYLKEPLNMYVYFHDQIITP 555 | CC EMBL: D50617; BAA09191.1; -; Genomic_DNA. | |
| Qy | 556 HFSPISHPANTVRVRLQRDIDVDSADMCLYALDEITDGFAVIGIYEATEDAYF 615 | CC DR SGD: S28-01; SCER-S28-01-001921-MONOMER; -. | |
| Db | 535 HFGPISHCANTRVRLQRDIDVDSADMCLYALDEITDGFAVIGIYEATEDAYF 594 | CC DR LinhHub: PA3553; -. | |
| Qy | 616 TARDTFSMMQIGCESERVMTLGRADIVKFARCOEDANSSEGTQFQFQIQQ 675 | CC DR Go: GO:005890; C: plasma membrane; IMP. | |
| Db | 595 MARDDFAMQLQIGESESRVMTLGRADIVKFARCOEDANSSEGTQFQFQIQQ 643 | CC DR Go: GO:0015084; F: di-, tri-valent inorganic cation transport; IDA. | |
| Qy | 676 QQQQQAPPPIPITSPINSTLNLSGTSTGGVGGVGGINGPNTGNTNTNTG 735 | CC DR InterPro: IPR02323; MG2+-transpCorA. | |
| Db | 644 -----GPALTQIN-LANLQARDINA-----SHKNNNSSTTPNRY 678 | CC DR Complete proteome; Magnesium; Membrane; Transport. | |
| Qy | 616 TARDTFSMMQIGCESERVMTLGRADIVKFARCOEDANSSEGTQFQFQIQQ 675 | CC DR FT TRANSMEM 742 762 | |
| Qy | 736 SPSPPOOOQQRGITNRKSPIDPARRADIALYKGTDHJLTFQNLAYERKFPSRSHEN 795 | CC DR FT TRANSMEM 772 792 | |
| Db | 679 APITSQ-----PRGDIALYLGTDHJLTFQNLAYERKFPSRSHEN 720 | CC SQ SEQUENCE 858 AA; 96684 MW; TAAFC6086F0618 CRC64; | |
| Qy | 796 YLAQLQVESFSNNNKITEMSKITLGMVTPNVLVTRGKVNPKGE-NSSIAWWFGI 855 | Query Match 34.6%; Score 1666; DB 1; Length 858; | |

| | | | |
|-----------------------|---|-----------------------------------|--------------|
| Best Local Similarity | 40.9% | Pred. | No. 3.1e-86; |
| Matches | 401; | Conservative | 126; |
| Indels | 221; | Gaps | 35; |
| | | | |
| Oy | 39 DSEDBLELKSLKSHLSE:SE: ---KEKQQQHQHQTESNAKPLTRKGSSTIKKSKNLTDKD 93 | | |
| Db | 9 DSSDLPLRSKSYDTAASMKGKYPKLENTRQ---YSDAQOPTRHE---ALAKVDETDKDS 62 | | |
| Oy | 94 RITNPMSLGGDDTINSKHNRY---AMS-SLRDFYLKONTDNSTNNTHALIPIP 148 | | |
| Db | 63 R--HRPSSNGB---NSGSGVYERNTNSTSRDGF--EGAAEAEATYKQLASFL- 114 | | |
| Oy | 149 IPIPTPIITNAKS-R-----RKSOLENLPLIKKK- ---TYGRNNSNFFND 191 | | |
| Db | 115 -----ISSNARPSRLAKSETHQ-QINVESTSPLKNAELRGHTPALGTSNNRCNL 168 | | |
| Oy | 192 LVSPTKTKTN-----DSEDTNT-----STTAHMLGLGATTL 226 | | |
| Db | 169 EASSARTTSARKASLVSIAFETSASEGHTPKQAKURRTYSTITHS- SYNPTL 226 | | |
| Oy | 227 GVGTTGTTATAATATAAG-----RPSRSTDSEADSHASRSQETEEDVCFPMVG- 276 | | |
| Db | 227 ----UUTRASOKSDMGNDRTRKPLMRDSEVFSEIT-SQASRDQETEEDVCFPMVL 280 | | |
| Oy | 277 DILRYNGIDPIDEIFIREREAYLQKQMLRIDEFQNLSSNNNTGASHH---P 333 | | |
| Db | 281 LITRNGYDPELEYA-----QISNAENNLNLAMHQHQSRT 318 | | |
| Oy | 334 YHHHSNNNNKNGDGSSSWALKYTPKPNLKLTKLSRFEPTHESSSEELYLKTKQ 393 | | |
| Db | 319 YHHTDQTCPTNSASTSOSS-AAKTPE---ISR---TLEKNCSTNEYAVSENNES 368 | | |
| Oy | 394 PPKYDOLISLTSSTSSTGGSGQVRPGARTDGNG-----SPD 437 | | |
| Db | 369 --VREDDKPDL-----HEDVTPFGNKTEGEKREGNDSSYRAYTTLQNTYQPS 415 | | |
| Oy | 438 RPSLPHSERBETHAPDPSLVSPGSVRLPLNGBETTMWDOCTCPSEMMOMLAKGJ 497 | | |
| Db | 416 RSPFRSESDTVAASDPSLSSGQYTFELKFQGDFTTWLDOCTPDEMRCKIAKNTG 475 | | |
| Oy | 498 HPLTAEDTRMOTBTKYELPKSYTYCPTFEADKBSDELYEPINVYIVFDGILTHPF 557 | | |
| Db | 476 HPLTAEDTRMOTBTKYELPKSYTYCPTFEADKBSDELYEPINVYIVFDGILTHPF 535 | | |
| Oy | 558 SPISHPANVRVRQRQLADYDVSADMWCYALIDEIDGPAFTHGLEYEADIEAVFTA 617 | | |
| Db | 536 DPISCANVRVRQRQLDYSVSNDWLYALIDDITDSTAPVQISYEADSIDSUVF 595 | | |
| Oy | 618 RUDPFSMLQIGRSRKVTLKLGRADVKMFKAKRQCEPANSSEYYORQ---FNL 674 | | |
| Db | 596 RUDPFAAMLQIGRSRKVTLKLGRADVKMFKAKRQCEPDTNGLIGEVLKJSQNMNL 655 | | |
| Oy | 675 QQQQQAPPPNPWITSPINSTNLSGTTGGVGVGINFGENPGCQNTNTNTTT 734 | | |
| Db | 656 QREQB-----NVQ-----NNSN----- 668 | | |
| Oy | 735 GSPSPPOOOQHGNTNKSPPDARPADTIALYLDQDIIHMTMFONLLAYEKIFSRSHS 794 | | |
| Db | 669 -----NQISLSN-SYQTSOPRGDIALYQDIIHMTMFONLLAYEKIFSRSHA 718 | | |
| Oy | 795 NYLAQLOQESTNSNKT-TEMFSKTIGLILVULNLFGRANTVPGEGTNLSWTFG 854 | | |
| Db | 719 NYLAQLOQESTNSNKT-TEMFSKTIGLILVULNLFGRANTVPGEGTNLSWTFG 777 | | |
| Oy | 855 IVGVLIIFTIGSFAQOMWLKLNST---ECONGNRP1-----PNHSSRSI 900 | | |
| Db | 778 ILGULLLAVSWPLASWIKKDPATNEAAGGSAKSIVSSLPKDRKPKRNDSKGN 837 | | |
| Oy | 901 RSLGLKKGKGNKSISPPNPK 921 | | |
| Db | 838 ARVGYTR---KSTVSLPSRY 854 | | |
| RESULT | 7 | | |
| Q7SAU1 | ASHGO | PRELIMINARY: | |
| ID | Q7SAU1 | ASHGO | PRT: 799 AA. |
| AC | Q7SAU1 | | |
| DT | 05-JUL-2004 | integrated into UniProtKB/TrEMBL. | |
| DT | 05-JUL-2004 | sequence version 1. | |
| DT | 07-FEB-2006 | entry version 16. | |

| | | | |
|-----------------------|--|------------------|----------------------|
| DE | ADL171Wp. | | |
| GN | OrderedLocusNames=ADL171W; Ashbya gossypii (Yeast) (Bremothecium gossypii). | | |
| OS | Eukaryota; Fungi; Ascomycota; Saccharomyctaceae; Saccharomyctes; | | |
| OC | Saccharomyctaceae; Saccharomyctes; Saccharomyctaceae; Bremothecium. | | |
| NCBI_TaxID | 33169; | | |
| RN | [1] | | |
| RP | NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA]. | | |
| RA | STRAIN=ATCC 10855; DOI=10.1126/science.1095781; | | |
| RA | Dietrich S., Voegeli S., Brachat S., Lerech A., Gates K., Steiner S., | | |
| RA | Mohr C., Poehlmann R., Luedi P., Choi S., Wing R.A., Platier A., | | |
| RA | Gaffney T.D., Philippson P.; | | |
| RT | *The Ashbya gossypii genome as a tool for mapping the ancient | | |
| RT | Saccharomyces cerevisiae genome.*; | | |
| RL | Science 304:304-307(2004). | | |
| CC | Copyright by the UniProt Consortium, see http://www.uniprot.org/terms | | |
| CC | Distributed under the Creative Commons Attribution-NoDerivs License | | |
| CC | CC | | |
| EMBL | A016817; AA551749.1; -; Genomic_DNA. | | |
| AGD | ADL171W; | | |
| DR | GO:0016030; C:membrane; IBA. | | |
| DR | GO:0046873; P:metal ion transporter activity; IBA. | | |
| DR | InterPro: IPR001003; Mg2+ transport; IBA. | | |
| DR | InterPro: IPR005253; Mg2+_transportCorA. | | |
| PFam | PF01544; CorA_1. | | |
| KW | Complete proteome; | | |
| SQ | SEQUENCE 799 AA; 89221 MW; DC04P4A752AB147A CRC644; | | |
| Query Match | 34.18; | Score 1641.5; | DB 2; Length 799; |
| Best Local Similarity | 42.8% | Pred. 7.1e-246; | Indels 161; Gaps 24; |
| Matches | 383; | Conservative 10; | Mismatches 246; |
| Qy | 62 QHHOBITSDNAKPLTRKSGSSIKKSNNLTDKDRITNPMSLGSQDDTINSGHKNRMVSS 121 | | |
| Db | 29 EDHANDLYDQHQPVALRKOSVADDSGPGRNVLNP---RSSEKQSRBENSQARLG 85 | | |
| Qy | 122 LRKYDYLKDPTDINSNRATH LAIPIPIPITPLTMANKSRKQLQENLPPLIKRTIG 181 | | |
| Db | 86 IRSKLVYDSEODYP---EHYR----PIPRATVSKSAGWP----TATASQTL- | | |
| Qy | 182 RNNNNNPENTDVSPMTOMKTNDSEDITNTSTPAHMKLGIGATTG-VGTG----TTAT 235 | | |
| Db | 132 ---YGPYPERA-----DTEKAGSCTGGGKLSKRSGRTSITVDGLRKMSRAQ 178 | | |
| Qy | 236 ATATAAG----RPRPSRTSDEADHARSQETEEDVCPM 274 | | |
| Db | 179 SETISOGNHLPSRTSIEKSQDSLUNTRNRRASED-DSNARSQETEEDVCPM 237 | | |
| Qy | 275 -VGDHIRVNGDIDDEDETFREBREBEAYLQKQMIAKNLTIDFPPNLSKNNTSGAARHP 333 | | |
| Db | 238 QPETHRTINGDIDDELETAQE-----IR-----ANLTANINGVSKEDD 278 | | |
| Qy | 334 YHHHSNNNNKNGDGSSMAALKTPKNNLTKLRTKPLKTTKQ----- 393 | | |
| Db | 279 VDMKCSCSPOSSTGTSTSISSAKMKTQPGNPKQDSNNETAPGFSGRNKVY--ECEY 335 | | |
| Qy | 394 PPYYDQDLSUTSSTSSTSGSGQYKFGEGARI SDGNGSLPDRSLFHSESBTIHP 453 | | |
| Db | 336 LP--MNDLSSTANEASYDMDGNIPADNYK-----PDRPSFCSAQGTVHP 384 | | |
| Qy | 454 DIPSLVSPGQSVRDLPRNGETMWLDCCTPDSKEMMLAKAFGIPILTAEDIHQETREK 513 | | |
| Db | 385 DIPSLVSPGQSVRDLPRNGETMWLDCCTPDSKEMMLAKAFGIPILTAEDIHQETREK 444 | | |
| Qy | 514 VEIKSYKPCPTEADKSESYLIPINNVTFPFDGILTHPESIPHARMVBRVQL 573 | | |
| Db | 445 VELFKSYKPCPTEADKSESYLIPINNVTFPFDGILTHPESIPHARMVBRVQL 504 | | |
| Qy | 574 RDYDVSAWICLALIDEITDGAFTGIEYADEIADEDVTFDTPSFMLORIGSRR 633 | | |
| Db | 505 RDYDVSAWICLALIDEITDGAFTGIEYADEIADEDVTFDTPSFMLORIGSRR 564 | | |
| Qy | 634 RKTPLRLSLGKADYTKPARKCQEANSSGGYTORQNUQQQQQAPPPIPITSP 693 | | |
| Db | 565 RKTPLRLSLGKADYTKPARKCQEANSSGGYTORQNUQQQQQAPPPIPITSP 602 | | |
| Qy | 694 INSTLNLSLGSITGSTGCGVGCGINFGPNTGNNNTNTGSPSPQQQQQHGJTKNSP 753 | | |

| | | | | | | | | | | |
|---|--|--------------|--|--|------|--------|------|------|-----|--|
| Matches | 310; | Conservative | 100; | Mismatches | 247; | Indels | 220; | Gaps | 22; | |
| Oy | 41 EDELEK-KSELESEVYSEKQHQEIT-----SDNAK 73 | RC | STRAIN=972 h-; | | | | | | | |
| | : : : : : : : : : : : : : : : : | RX | MEDLINE=20089027; PubMed=10620777; | | | | | | | |
| Db | 91 ESEIEHLHESENVPQPKKKRKRAKSRKIPATOSSEADMTHVCPGSGSDRK 150 | RX | DOI=10.1002/(SICI)1097-0051(20000115)16:1<71::AID-YEAS053>3.0.CO;2-5; | | | | | | | |
| | : : : : : : : : : : : : : : : : | RA | Haitawa Y., Yamazaki J., Yamamoto S., Tanaka T., Kusuda N., Jinno K., | | | | | | | |
| Db | 74 PIUTRS-GSSIK-KISNLTJDKURINPNSL-SGGDTTNSGHKNMNY-----MSSRK 124 | RA | Sakai M., Aoki K., Ogura K., Kubuchi H., Zhang M.Q., Yangide M. | | | | | | | |
| | : : : : : : : : : : : : : : : : | RA | * A 38 kb segment containing the cdc2 gene from the left arm of fission yeast chromosome I: sequence analysis and characterization of the genomic DNA and cDNAs encoded on the segment.; | | | | | | | |
| Db | 151 POSNKKHRGRVHSKPSKTLVEPGDPIKALSAISGTS-SQASH-YNEGRGSFPTJGT 208 | RT | Year 16:71-80 (2000). | | | | | | | |
| | : : : : : : : : : : : : : : : : | RL | Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms | | | | | | | |
| Oy | 125 DEYL-KDNTDINSTNNTHLAIPIPIPPIPPNLSKRSKQLNLENPLPKTKTIGRN 183 | RN [2] | CC | Distributed under the Creative Commons Attribution License | | | | | | |
| | : : : : : : : : : : : : : : : : | RP | RPNUCLEOTIDE SEQUENCE. | | | | | | | |
| Db | 209 HYTHLSKSRKSSSDSYSSVMSMSKNDNSDLSSSSPVGSDDNLNPL-IDKTYLTV 266 | RC | STRAIN=972 h-i | | | | | | | |
| | : : : : : : : : : : : : : : : : | RA | Director-General of Biotechnology Center, Kusuda N., Machida M., Director-General of GenBank/DBJ databases. | | | | | | | |
| Db | 184 NSNNPENDLVSMPMTKMTNSEDINTNTSTPANTNMKGIGATLGV-GCTTTATATAAA 242 | RA | Submitted (MAy-1997) to the EMBL/GenBank/DBJ databases. | | | | | | | |
| | : : : : : : : : : : : : : : : : | RL | CC | Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms | | | | | | |
| Oy | 267 EDPSFAVHRPRPSATKSCSAAVDCPHTIPKPPVQSDTDLTELPSTSTQFSDETFtQP 326 | CC | CC | CC | | | | | | |
| | : : : : : : : : : : : : : : : : | CC | Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms | | | | | | | |
| Oy | 243 GRRPSRSSI---DSEADSHARSQSQTTEEDVCPMPGDHIRUNGDFDDEDEFIRERE 298 | CC | CC | CC | | | | | | |
| | : : : : : : : : : : : : : : : : | CC | Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms | | | | | | | |
| Db | 327 RLKPHSSSTANDEBREDSQDVENDTRVVEDCVCPPQEEHVNKGDFDELNFNEE-----384 | DR | EMBL: AB004539; BA21448.1; -; Genomic_DNA. | | | | | | | |
| | : : : : : : : : : : : : : : : : | DR | GO: GO:0016020; C:membrane; IEA. | | | | | | | |
| Oy | 299 EAYLKQMTANKLRIDEFQNLKSNTGASHR-----PYHHHSNNNKKANGGCGSSS 353 | DR | CO: CO:0046872; P:anion transporter activity; IEA. | | | | | | | |
| | : : : : : : : : : : : : : : : : | DR | GO: GO:0030001; P:metal ion transport; IEA. | | | | | | | |
| Db | 385 ---LQRQ-----RQNDHFRSQRYSKTCRPEFHNDLSPHDND-----420 | DR | InterPro: IPR02523; Mg2+-transp;CORA. | | | | | | | |
| | : : : : : : : : : : : : : : : : | PFam: | Pf01544; CORA_1. | | | | | | | |
| Oy | 354 MAALKYPTNLKKTLSRFEPTHEINSSSSEPTIYQKTDOLSLTSSTS-TSG 413 | DR | SEQUENCE 1 | | | | | | | |
| | : : : : : : : : : : : : : : : : | FT | 56728 MW; | | | | | | | |
| Db | 421 -----PSSL-----HSNNAAKAEPVPLRSS--YSSGGRSVTMSMSD-----456 | FT | BBC3715EB62D2C4C CRC64; | | | | | | | |
| | : : : : : : : : : : : : : : : : | SQ | SEQUENCE 499 AA; | | | | | | | |
| Oy | 414 SGSGQVKFGARISDGINGSLPDRFLSESEETHADPILWSPG-QSVRLPFLNG 472 | FT | Score 1141.5; DB 2; | | | | | | | |
| | : : : : : : : : : : : : : : : : | FT | Length 499; | | | | | | | |
| Db | 457 -----PTRYFFESSENQNETHASYLDLDDGATSFRLSFCPE 494 | FT | Best Local Similarity 40.96; Freq. No. 1.4e-56; | | | | | | | |
| | : : : : : : : : : : : : : : : : | FT | Indels 143; Gaps 13; | | | | | | | |
| Oy | 473 EETWWILDCTCPTDSEMCKMLAKAFGIIHPLTAEDIRMOETREKVELPKSYYFVCFHTEADK 532 | Qy | 237: 232: TTAAAGGRRPSSSI----DSEADSHARSQSQTTEEDVCPMPGDHIRVN 282 | | | | | | | |
| | : : : : : : : : : : : : : : : : | Qy | 7 TKTAQPSDEFTVQPRLKSTHSSTADNEEVDYDQYDENTVVEEDVCPMPQESHSYWK 66 | | | | | | | |
| Db | 495 KGWWNCLCDLPDTEMLRVLKSFISHLTETDVRQBARKEKELFSTYVCFERSEFQDP 554 | Qy | 238 QDPPDEDFPFREREAYAQKOMIAKNLRIDEFQNLKNNTGASHR-----PYFRH 337 | | | | | | | |
| | : : : : : : : : : : : : : : : : | Qy | 67 GIDPBDLNPAEE-----LQKQ-----RONTDPRHSQYSTCPPEFH 105 | | | | | | | |
| Oy | 533 ESDYDLEPINVIVWPAGLITIFPSISPARVQIQLDYDVIDAWUICYALDI 592 | Qy | 338 SNNNKUNGCGGSSMSMAALKYPTNLKKTLSRFTHEINSSSEBEITLTKOQQPPK 397 | | | | | | | |
| | : : : : : : : : : : : : : : : : | Qy | 106 WNDLSPHDND-----PSSSL-----HSNNAAKAEPVPLRSS--RY 139 | | | | | | | |
| Db | 555 ELANTLEPLAIVVIVWPAGLITIFPSISPARVQIQLDYDVIDAWUICYALIDI 614 | Qy | 339 YDQQLSITSTPSGSSCQWFGGARISDGINGSUPRASESSEBTETHADPIS 457 | | | | | | | |
| | : : : : : : : : : : : : : : : : | Qy | 140 SGGRSITSMDS-----PSSSL-----PYRFSFSNNEONETHASVLSUD 174 | | | | | | | |
| Oy | 593 TDGPAPVHGLEYEADATEVADTAFTARDPSSMLQIGRESRKVMTLRLSKGADVKM 652 | Qy | 458 LVSPE-QEVROLFNRGEETWMDCTPDSSEMMLAKAFGIIHPLTAEDIRMOETREKVEL 516 | | | | | | | |
| | : : : : : : : : : : : : : : : : | Qy | 175 LLDGAGATPSRLFPCPGVNMFLCDPDTDEMVLSKAStHPLTEDRVQEAEKVEL 234 | | | | | | | |
| Db | 615 TDAFPVLRGLETEAEDSVLUGRSSEDDSDMLBRGECKATGMFLRKGADVKM 674 | Qy | 517 FKSTYFICPHFPAKSEDSLIPBENYIVWFDGLLTHPSPISHANRARRVQDLY 576 | | | | | | | |
| | : : : : : : : : : : : : : : : : | Qy | 235 FGSTYFICPHFPAKSEDSLIPBENYIVWFDGLLTHPSPISHANRARRVQDLY 294 | | | | | | | |
| Oy | 653 FAKRQCEAANSSGGYQYQINLQ0000QAPPNPNLITPSINLNLSTGTVGG 712 | Qy | 577 VDVSADMVCALEIDTDFGAPVHGLEYEADATEVADTAFTARDPSSMLQIGRESRK 636 | | | | | | | |
| | : : : : : : : : : : : : : : : : | Qy | 205 VHVNSDWLCAKUJIDTDAVPLRGTEETEIDSVLGSESDSEMURIGCRKT 354 | | | | | | | |
| Db | 675 LAKRNEKWTIA----- | Qy | 637 MTLARLUSGADYVTKMFACIKQBANSSSGGYQYQINLQ0000QAPPNPNLITPS 696 | | | | | | | |
| | : : : : : : : : : : : : : : : : | Qy | 355 NGMPRLYGRADVTKMFLAKCNKWTIA----- | | | | | | | |
| Db | 687 -----PFG-----EGLIGDQ 699 | Qy | 355 NGMPRLYGRADVTKMFLAKCNKWTIA----- | | | | | | | |
| | : : : : : : : : : : : : : : : : | Qy | 382 | | | | | | | |
| Oy | 773 DHITTMFONLAYERKIFSRSHSYLAQLOVESFSNMRKTFITLGMVLPANYT 832 | Qy | 697 TLNLNSLGTSTCGGVGGINGFNPPIPGNNTNTGSPSPRQQQOHGITNKSPPTP 756 | | | | | | | |
| | : : : : : : : : : : : : : : : : | Qy | 383 | | | | | | | |
| Db | 700 DILVMTMSNLSEOFENLRSRTHSYLAQLOTNSEFNTSITNGALKTGTLVPMNVT 759 | Qy | 383 | | | | | | | |
| | : : : : : : : : : : : : : : : : | Qy | 383 | | | | | | | |
| Oy | 833 GLGCHAVNVRPREGGTMGWPGFIVGVL-SIIGSPT 868 | Qy | 383 | | | | | | | |
| | : : : : : : : : : : : : : : : : | Qy | 383 | | | | | | | |
| Db | 760 GLFGAVNPVPGDNTNLLAWFFGIVGLGTAIWII 796 | Qy | 383 | | | | | | | |
| | : : : : : : : : : : : : : : : : | Qy | 383 | | | | | | | |
| RESULT 11 | | Qy | 383 | | | | | | | |
| Q1LWDB_SCHPO | PRELIMINARY; | PRP: | 499 AA. | | | | | | | |
| ID | Q1LWDB_SCHPO | PRIMINARY; | PRP: | 499 AA. | | | | | | |
| AC | 05-07-2004, integrated into UniProtKB/TrEMBL. | | | | | | | | | |
| DT | 05-JUL-2004, sequence version 1. | | | | | | | | | |
| DT | 07-FEB-2005, entry version 8. | | | | | | | | | |
| Probable membrane protein YOL130w (Fragment). | | | | | | | | | | |
| GN | Name=Pi66; | | | | | | | | | |
| OS | Schizosaccharomyces pombe (Fission yeast). | | | | | | | | | |
| OC | Eukaryota; Fungi; Ascomycota; Schizosaccharomyces pombe. | | | | | | | | | |
| OC | Schizosaccharomyces; Schizosaccharomyces pombe. | | | | | | | | | |
| NCBI TaxID=4836; | | | | | | | | | | |
| RN | [1] | | | | | | | | | |
| RP | NUCLEOTIDE SEQUENCE. | | | | | | | | | |

| | | | |
|-----------------------|---|----|--|
| DR | GO:0030001; Pimetal ion transport; IFA. | RA | Lui A., Ma L.-J., Mabbitt R., MacDonald J., MacLean C., Major J., Manning J., Maru K., Matthee C., Mauceli E., McCarthy M., McDonough S., McHugh T., Meldrum J., Menetsky K., Nesirov J., Mikhael A., Mihova T., Mileng V., Moru K., Mozes J., Mulrain L., Munson G., Nayyeri J., Newes C., Nguyen C., Nguyen N., Nguyen T., Nicols R., Nielsen C., Norbu C., Norbu N., O'Donnell P., Okawa O., O'Leary S., Omotochou b., O'Neill K., Osman S., Parker S., Perrin D., Phunkhang P., Picani B., Purcell S., Rachupka T., Ramamurthy U., Rameau R., Ray V., Raymond C., Reita R., Richardson S., Rodriguez J., Rogers J., Rogov P., Rutman M., Schupbach R., Seaman J., Sharpe T., Sounez C., Sheridan J., Shera N., Shi J., Smirnov S., Smith C., Sounez C., Spencer B., Stalker J., Stange-Thiermann N., Stavropoulos S., Stenson K., Stone C., Stone S., Stubbs M., Talama J., Tchuinga P., Tenzing P., Tesfaye S., Theodore J., Thounthang Y., Topham K., Towey S., Tsamli T., Tsomo N., Vallee D., Vassiliev H., Venkataraman V.S., Vinson J., Vo A., Wade C., Wang S., Wangchuk T., Wangi T., Whittaker C., Wilkinson J., Wu J., Wyman D., Yadav S., Yang S., Yang X., Yeager S., Yee B., Young G., Zainoun J., Zembeck L., Zimmer A., Zody M., Lander E.S.; |
| DR | InterPro; IPR01356; Homeobox. | RA | Romanblad-Toth K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O., Lindblad-Toth K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O., |
| DR | InterPro; IPR01253; Mg++_transportCora. | RA | |
| DR | Pf01544; CorA; 1. | RA | |
| DR | PROSITE; PS00027; HOMEOXO_1; UNKNOWN_1. | RA | |
| KW | Complete proteome. | RA | |
| SEQUENCE | 460 AA; | RA | |
| SQ | 52281 MW; | RA | |
| Score | 1044; | RA | |
| Length | 460; | RA | |
| Best Local Similarity | 21.7%; | RA | |
| Matches | 42.5%; Pred. No. 3.8-51; | RA | |
| Matches | 202; Conservativeness 93; Indels 104; Gaps 3; | RA | |
| Oy | DREFLPHSESEFTIADPISLPSGOSVRLPFNGTETWLDCTCPDSEMOKMLAKAG 496 | RA | |
| Oy | 77 DIPIAFSTCECTTRANEFSSEIERDGCKAKELPAFKHGTMWLCYDPTDAENKLAKAG 136 | RA | |
| Db | 497 IHLTAEDTRMOTETKELYTFVCFRHTFADKSESEDYLEDPIVINYLVFDGLTH 556 | RA | |
| Oy | 137 IHPFLTYDQHREAKEVTEMNRYTFAFHAYBODVESDFMFMVAFYLCLVPEGVISH 196 | RA | |
| Db | 557 FSPISHPANVERRVERVQLRDYDVSDADMUYCALIDEDETDGAPVTHGIBYADEIADEAFT 616 | RA | |
| Oy | 197 YSPVPICATVRKRIQKQLDHYTPDWICAVIDDTSDAPVTRIEQDGELEVEQTVD 256 | RA | |
| Oy | 617 ARDPODESSLQHIGTNSRUKVATLRLSGRADYTFKPARKCQBEANSSGGYQRQNLQO 676 | RA | |
| Db | 257 ARDGFNKLRLIGKARAKUTSMRLLSGRADYVGMETKVSF----- 299 | RA | |
| Oy | 677 QQQQAPPPIPNTSPINSLNLSLGTSTGGVGYGGINFGPNPFGGNNTNTNTCS 736 | RA | |
| Db | 300 ----- 299 | RA | |
| Oy | 737 PSPRQOCQGHGTMKSPPIPDARPADIALYLGDIOPHITMFQNLLAYEKIFSRSHSY 796 | RA | |
| Db | 300 ----- GLSDRVPKGHTSILYLGDIOPHITMFQNLLAYEKIFSRSHSY 342 | RA | |
| Oy | 797 LAQLOQVEPNSPNKKTENFSKTTLGMILPLVNLYCPLGMYTRVCEGST---NLGMP 852 | RA | |
| Db | 343 LSQLOVQSDIAHRYTDLGLKITYTGITLIPMNVTGLFAMVNRIGESTDENGPNF3MF 402 | RA | |
| Oy | 853 FGIVELPLITGTSPTIAQWNLKYLANSITCQNNNGRPFIENSSRSIRSISLGLK 907 | RA | |
| Db | 403 FGILGYLCLITVATUAKWLNAYKEKIRGN-----FRNSASDNTSRSMSKR 451 | RA | |
| RESULT 1 | 04 PHQ2_USTMA PRELIMINARY; PRT: 881 AA. | RA | |
| ID | OAPHQ2_USTMA | RA | |
| OAPHQ2_USTMA | PRELIMINARY; PRT: 881 AA. | RA | |
| AC | 04 PHQ2_USTMA | RA | |
| DT | 19-JUL-2005; integrated into UniProtKB/TREMBL. | RA | |
| DT | 19-JUL-2005; sequence version 1. | RA | |
| DT | 07-FEB-2006; entry version 4. | RA | |
| DS | Hypothetical protein. | RA | |
| GN | ORFName=UR0361.1; | RA | |
| CC | Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes; Ustilaginaceae; Ustilago. | RA | |
| OS | Ustilago maydis 521. | RA | |
| OC | Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago. | RA | |
| NCBI_TaxID | 1237631; [1] | RA | |
| RN | NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA]. | RA | |
| RC | STRAIN=21; | RA | |
| RA | Birren B.W., Nusbaum C., Abebe A., Abouelleil A., Adekoya E., Ait-Zahra M., Allen N., Allen T., An P., Anderson M., Anderson S., Arachchi H.M., Ambrosius J., Bachantash P., Baldwin J., Barry A., Bayul T., Blitshteyn B., Bloom T., Blive J., Bogoslavskiy L., Borowsky M., Bouhala B., Brunache A., Butler J., Calixte N., Calvo S.E., Camarata J., Campo K., Chang J., Cheshatsang Y., Collymore A., Condidine T., Cook A., Cooke P., Corum B., Citron M., Collymore A., David R., Dawoe T., Degray S., Dodge S., Dooley K., Cuomo C., Dorrie K., Dorris L., Duffey N., Dupes A., Ekins T., Engels R., Erickson J., Farina A., Ferreira P., Fischer H., Fitzgerald M., Foley K., Gage D., Galagan J.E., Gearin G., Gneire S., Gnirke A., Govertse A., Graham J., Grandbois E., Gyaltset K., Hafez N., Hagopian D., Hagus B., Hall J., Fletcher B., Heiter A., Higgins H., Honan T., Horn A., Houde N., Hughes L., Huime W., Husby E., Ilieev I., Jaffe D., Jones C., Kamal M., Kamysheva M., Karlsson E., Kells C., Kieser A., Kiser P., Kodira C., Kuloboks E., Lubutti K., Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T., Lindblad-Toth K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O., | | |
| RA | Score 1044; | RA | |
| DB | Length 460; | RA | |
| RA | Best Local Similarity 21.7%; | RA | |
| RA | Matches 42.5%; Pred. No. 3.8-51; | RA | |
| RA | Matches 202; Conservativeness 93; Indels 104; Gaps 3; | RA | |
| Oy | DREFLPHSESEFTIADPISLPSGOSVRLPFNGTETWLDCTCPDSEMOKMLAKAG 496 | RA | |
| RA | 77 DIPIAFSTCECTTRANEFSSEIERDGCKAKELPAFKHGTMWLCYDPTDAENKLAKAG 136 | RA | |
| Db | 497 IHLTAEDTRMOTETKELYTFVCFRHTFADKSESEDYLEDPIVINYLVFDGLTH 556 | RA | |
| Oy | 137 IHPFLTYDQHREAKEVTEMNRYTFAFHAYBODVESDFMFMVAFYLCLVPEGVISH 196 | RA | |
| Db | 557 FSPISHPANVERRVERVQLRDYDVSDADMUYCALIDEDETDGAPVTHGIBYADEIADEAFT 616 | RA | |
| Oy | 197 YSPVPICATVRKRIQKQLDHYTPDWICAVIDDTSDAPVTRIEQDGELEVEQTVD 256 | RA | |
| Oy | 617 ARDPODESSLQHIGTNSRUKVATLRLSGRADYTFKPARKCQBEANSSGGYQRQNLQO 676 | RA | |
| Db | 257 ARDGFNKLRLIGKARAKUTSMRLLSGRADYVGMETKVSF----- 299 | RA | |
| Oy | 677 QQQQAPPPIPNTSPINSLNLSLGTSTGGVGYGGINFGPNPFGGNNTNTCS 736 | RA | |
| Db | 300 ----- 299 | RA | |
| Oy | 737 PSPRQOCQGHGTMKSPPIPDARPADIALYLGDIOPHITMFQNLLAYEKIFSRSHSY 796 | RA | |
| Db | 300 ----- GLSDRVPKGHTSILYLGDIOPHITMFQNLLAYEKIFSRSHSY 342 | RA | |
| Oy | 797 LAQLOQVEPNSPNKKTENFSKTTLGMILPLVNLYCPLGMYTRVCEGST---NLGMP 852 | RA | |
| Db | 343 LSQLOVQSDIAHRYTDLGLKITYTGITLIPMNVTGLFAMVNRIGESTDENGPNF3MF 402 | RA | |
| Oy | 853 FGIVELPLITGTSPTIAQWNLKYLANSITCQNNNGRPFIENSSRSIRSISLGLK 907 | RA | |
| Db | 403 FGILGYLCLITVATUAKWLNAYKEKIRGN-----FRNSASDNTSRSMSKR 451 | RA | |
| RESULT 1 | 04 PHQ2_USTMA PRELIMINARY; PRT: 881 AA. | RA | |
| ID | OAPHQ2_USTMA | RA | |
| OAPHQ2_USTMA | PRELIMINARY; PRT: 881 AA. | RA | |
| AC | 04 PHQ2_USTMA | RA | |
| DT | 19-JUL-2005; integrated into UniProtKB/TREMBL. | RA | |
| DT | 19-JUL-2005; sequence version 1. | RA | |
| DT | 07-FEB-2006; entry version 4. | RA | |
| DS | Hypothetical protein. | RA | |
| GN | ORFName=UR0361.1; | RA | |
| CC | Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes; Ustilaginaceae; Ustilago. | RA | |
| OS | Ustilago maydis 521. | RA | |
| OC | Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago. | RA | |
| NCBI_TaxID | 1237631; [1] | RA | |
| RN | NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA]. | RA | |
| RC | STRAIN=21; | RA | |
| RA | Birren B.W., Nusbaum C., Abebe A., Abouelleil A., Adekoya E., Ait-Zahra M., Allen N., Allen T., An P., Anderson M., Anderson S., Arachchi H.M., Ambrosius J., Bachantash P., Baldwin J., Barry A., Bayul T., Blitshteyn B., Bloom T., Blive J., Bogoslavskiy L., Borowsky M., Bouhala B., Brunache A., Butler J., Calixte N., Calvo S.E., Camarata J., Campo K., Chang J., Cheshatsang Y., Collymore A., Condidine T., Cook A., Cooke P., Corum B., Citron M., Collymore A., David R., Dawoe T., Degray S., Dodge S., Dooley K., Cuomo C., Dorrie K., Dorris L., Duffey N., Dupes A., Ekins T., Engels R., Erickson J., Farina A., Ferreira P., Fischer H., Fitzgerald M., Foley K., Gage D., Galagan J.E., Gearin G., Gneire S., Gnirke A., Govertse A., Graham J., Grandbois E., Gyaltset K., Hafez N., Hagopian D., Hagus B., Hall J., Fletcher B., Heiter A., Higgins H., Honan T., Horn A., Houde N., Hughes L., Huime W., Husby E., Ilieev I., Jaffe D., Jones C., Kamal M., Kamysheva M., Karlsson E., Kells C., Kieser A., Kiser P., Kodira C., Kuloboks E., Lubutti K., Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T., Lindblad-Toth K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O., | | |
| RA | Score 1044; | RA | |
| DB | Length 460; | RA | |
| RA | Best Local Similarity 21.7%; | RA | |
| RA | Matches 42.5%; Pred. No. 3.8-51; | RA | |
| RA | Matches 202; Conservativeness 93; Indels 104; Gaps 3; | RA | |
| Oy | DREFLPHSESEFTIADPISLPSGOSVRLPFNGTETWLDCTCPDSEMOKMLAKAG 496 | RA | |
| RA | 77 DIPIAFSTCECTTRANEFSSEIERDGCKAKELPAFKHGTMWLCYDPTDAENKLAKAG 136 | RA | |
| Db | 497 IHLTAEDTRMOTETKELYTFVCFRHTFADKSESEDYLEDPIVINYLVFDGLTH 556 | RA | |
| Oy | 137 IHPFLTYDQHREAKEVTEMNRYTFAFHAYBODVESDFMFMVAFYLCLVPEGVISH 196 | RA | |
| Db | 557 FSPISHPANVERRVERVQLRDYDVSDADMUYCALIDEDETDGAPVTHGIBYADEIADEAFT 616 | RA | |
| Oy | 197 YSPVPICATVRKRIQKQLDHYTPDWICAVIDDTSDAPVTRIEQDGELEVEQTVD 256 | RA | |
| Oy | 617 ARDPODESSLQHIGTNSRUKVATLRLSGRADYTFKPARKCQBEANSSGGYQRQNLQO 676 | RA | |
| Db | 257 ARDGFNKLRLIGKARAKUTSMRLLSGRADYVGMETKVSF----- 299 | RA | |
| Oy | 677 QQQQAPPPIPNTSPINSLNLSLGTSTGGVGYGGINFGPNPFGGNNTNTCS 736 | RA | |
| Db | 300 ----- 299 | RA | |
| Oy | 737 PSPRQOCQGHGTMKSPPIPDARPADIALYLGDIOPHITMFQNLLAYEKIFSRSHSY 796 | RA | |
| Db | 300 ----- GLSDRVPKGHTSILYLGDIOPHITMFQNLLAYEKIFSRSHSY 342 | RA | |
| Oy | 797 LAQLOQVEPNSPNKKTENFSKTTLGMILPLVNLYCPLGMYTRVCEGST---NLGMP 852 | RA | |
| Db | 343 LSQLOVQSDIAHRYTDLGLKITYTGITLIPMNVTGLFAMVNRIGESTDENGPNF3MF 402 | RA | |
| Oy | 853 FGIVELPLITGTSPTIAQWNLKYLANSITCQNNNGRPFIENSSRSIRSISLGLK 907 | RA | |
| Db | 403 FGILGYLCLITVATUAKWLNAYKEKIRGN-----FRNSASDNTSRSMSKR 451 | RA | |
| RESULT 1 | 04 PHQ2_USTMA PRELIMINARY; PRT: 881 AA. | RA | |
| ID | OAPHQ2_USTMA | RA | |
| OAPHQ2_USTMA | PRELIMINARY; PRT: 881 AA. | RA | |
| AC | 04 PHQ2_USTMA | RA | |
| DT | 19-JUL-2005; integrated into UniProtKB/TREMBL. | RA | |
| DT | 19-JUL-2005; sequence version 1. | RA | |
| DT | 07-FEB-2006; entry version 4. | RA | |
| DS | Hypothetical protein. | RA | |
| GN | ORFName=UR0361.1; | RA | |
| CC | Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes; Ustilaginaceae; Ustilago. | RA | |
| OS | Ustilago maydis 521. | RA | |
| OC | Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago. | RA | |
| NCBI_TaxID | 1237631; [1] | RA | |
| RN | NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA]. | RA | |
| RC | STRAIN=21; | RA | |
| RA | Birren B.W., Nusbaum C., Abebe A., Abouelleil A., Adekoya E., Ait-Zahra M., Allen N., Allen T., An P., Anderson M., Anderson S., Arachchi H.M., Ambrosius J., Bachantash P., Baldwin J., Barry A., Bayul T., Blitshteyn B., Bloom T., Blive J., Bogoslavskiy L., Borowsky M., Bouhala B., Brunache A., Butler J., Calixte N., Calvo S.E., Camarata J., Campo K., Chang J., Cheshatsang Y., Collymore A., Condidine T., Cook A., Cooke P., Corum B., Citron M., Collymore A., David R., Dawoe T., Degray S., Dodge S., Dooley K., Cuomo C., Dorrie K., Dorris L., Duffey N., Dupes A., Ekins T., Engels R., Erickson J., Farina A., Ferreira P., Fischer H., Fitzgerald M., Foley K., Gage D., Galagan J.E., Gearin G., Gneire S., Gnirke A., Govertse A., Graham J., Grandbois E., Gyaltset K., Hafez N., Hagopian D., Hagus B., Hall J., Fletcher B., Heiter A., Higgins H., Honan T., Horn A., Houde N., Hughes L., Huime W., Husby E., Ilieev I., Jaffe D., Jones C., Kamal M., Kamysheva M., Karlsson E., Kells C., Kieser A., Kiser P., Kodira C., Kuloboks E., Lubutti K., Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T., Lindblad-Toth K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O., | | |
| RA | Score 1044; | RA | |
| DB | Length 460; | RA | |
| RA | Best Local Similarity 21.7%; | RA | |
| RA | Matches 42.5%; Pred. No. 3.8-51; | RA | |
| RA | Matches 202; Conservativeness 93; Indels 104; Gaps 3; | RA | |
| Oy | DREFLPHSESEFTIADPISLPSGOSVRLPFNGTETWLDCTCPDSEMOKMLAKAG 496 | RA | |
| RA | 77 DIPIAFSTCECTTRANEFSSEIERDGCKAKELPAFKHGTMWLCYDPTDAENKLAKAG 136 | RA | |
| Db | 497 IHLTAEDTRMOTETKELYTFVCFRHTFADKSESEDYLEDPIVINYLVFDGLTH 556 | RA | |
| Oy | 137 IHPFLTYDQHREAKEVTEMNRYTFAFHAYBODVESDFMFMVAFYLCLVPEGVISH 196 | RA | |
| Db | 557 FSPISHPANVERRVERVQLRDYDVSDADMUYCALIDEDETDGAPVTHGIBYADEIADEAFT 616 | RA | |
| Oy | 197 YSPVPICATVRKRIQKQLDHYTPDWICAVIDDTSDAPVTRIEQDGELEVEQTVD 256 | RA | |
| Oy | 617 ARDPODESSLQHIGTNSRUKVATLRLSGRADYTFKPARKCQBEANSSGGYQRQNLQO 676 | RA | |
| Db | 257 ARDGFNKLRLIGKARAKUTSMRLLSGRADYVGMETKVSF----- 299 | RA | |
| Oy | 677 QQQQAPPPIPNTSPINSLNLSLGTSTGGVGYGGINFGPNPFGGNNTNTCS 736 | RA | |
| Db | 300 ----- 299 | RA | |
| Oy | 737 PSPRQOCQGHGTMKSPPIPDARPADIALYLGDIOPHITMFQNLLAYEKIFSRSHSY 796 | RA | |
| Db | 300 ----- GLSDRVPKGHTSILYLGDIOPHITMFQNLLAYEKIFSRSHSY 342 | RA | |
| Oy | 797 LAQLOQVEPNSPNKKTENFSKTTLGMILPLVNLYCPLGMYTRVCEGST---NLGMP 852 | RA | |
| Db | 343 LSQLOVQSDIAHRYTDLGLKITYTGITLIPMNVTGLFAMVNRIGESTDENGPNF3MF 402 | RA | |
| Oy | 853 FGIVELPLITGTSPTIAQWNLKYLANSITCQNNNGRPFIENSSRSIRSISLGLK 907 | RA | |
| Db | 403 FGILGYLCLITVATUAKWLNAYKEKIRGN-----FRNSASDNTSRSMSKR 451 | RA | |
| RESULT 1 | 04 PHQ2_USTMA PRELIMINARY; PRT: 881 AA. | RA | |
| ID | OAPHQ2_USTMA | RA | |
| OAPHQ2_USTMA | PRELIMINARY; PRT: 881 AA. | RA | |
| AC | 04 PHQ2_USTMA | RA | |
| DT | 19-JUL-2005; integrated into UniProtKB/TREMBL. | RA | |
| DT | 19-JUL-2005; sequence version 1. | RA | |
| DT | 07-FEB-2006; entry version 4. | RA | |
| DS | Hypothetical protein. | RA | |
| GN | ORFName=UR0361.1; | RA | |
| CC | Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes; Ustilaginaceae; Ustilago. | RA | |
| OS | Ustilago maydis 521. | RA | |
| OC | Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago. | RA | |
| NCBI_TaxID | 1237631; [1] | RA | |
| RN | NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA]. | RA | |
| RC | STRAIN=21; | RA | |
| RA | Birren B.W., Nusbaum C., Abebe A., Abouelleil A., Adekoya E., Ait-Zahra M., Allen N., Allen T., An P., Anderson M., Anderson S., Arachchi H.M., Ambrosius J., Bachantash P., Baldwin J., Barry A., Bayul T., Blitshteyn B., Bloom T., Blive J., Bogoslavskiy L., Borowsky M., Bouhala B., Brunache A., Butler J., Calixte N., Calvo S.E., Camarata J., Campo K., Chang J., Cheshatsang Y., Collymore A., Condidine T., Cook A., Cooke P., Corum B., Citron M., Collymore A., David R., Dawoe T., Degray S., Dodge S., Dooley K., Cuomo C., Dorrie K., Dorris L., Duffey N., Dupes A., Ekins T., Engels R., Erickson J., Farina A., Ferreira P., Fischer H., Fitzgerald M., Foley K., Gage D., Galagan J.E., Gearin G., Gneire S., Gnirke A., Govertse A., Graham J., Grandbois E., Gyaltset K., Hafez N., Hagopian D., Hagus B., Hall J., Fletcher B., Heiter A., Higgins H., Honan T., Horn A., Houde N., Hughes L., Huime W., Husby E., Ilieev I., Jaffe D., Jones C., Kamal M., Kamysheva M., Karlsson E., Kells C., Kieser A., Kiser P., Kodira C., Kuloboks E., Lubutti K., Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T., Lindblad-Toth K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O., | | |
| RA | Score 1044; | RA | |
| DB | Length 460; | RA | |
| RA | Best Local Similarity 21.7%; | RA | |
| RA | Matches 42.5%; Pred. No. 3.8-51; | RA | |
| RA | Matches 202; Conservativeness 93; Indels 104; Gaps 3; | RA | |
| Oy | DREFLPHSESEFTIADPISLPSGOSVRLPFNGTETWLDCTCPDSEMOKMLAKAG 496 | RA | |
| RA | 77 DIPIAFSTCECTTRANEFSSEIERDGCKAKELPAFKHGTMWLCYDPTDAENKLAKAG 136 | RA | |
| Db | 497 IHLTAEDTRMOTETKELYTFVCFRHTFADKSESEDYLEDPIVINYLVFDGLTH 556 | RA | |
| Oy | 137 IHPFLTYDQHREAKEVTEMNRYTFAFHAYBODVESDFMFMVAFYLCLVPEGVISH 196 | RA | |
| Db | 557 FSPISHPANVERRVERVQLRDYDVSDADMUYCALIDEDETDGAPVTHGIBYADEIADEAFT 616 | RA | |
| Oy | 197 YSPVPICATVRKRIQKQLDHYTPDWICAVIDDTSDAPVTRIEQDGELEVEQTVD 256 | RA | |
| Oy | 617 ARDPODESSLQHIGTNSRUKVATLRLSGRADYTFKPARKCQBEANSSGGYQRQNLQO 676 | RA | |
| Db | 257 ARDGFNKLRLIGKARAKUTSMRLLSGRADYVGMETKVSF----- 299 | RA | |
| Oy | 677 QQQQAPPPIPNTSPINSLNLSLGTSTGGVGYGGINFGPNPFGGNNTNTCS 736 | RA | |
| Db | 300 ----- 299 | RA | |
| Oy | 737 PSPRQOCQGHGTMKSPPIPDARPADIALYLGDIOPHITMFQNLLAYEKIFSRSHSY 796 | RA | |
| Db | 300 ----- GLSDRVPKGHTSILYLGDIOPHITMFQNLLAYEKIFSRSHSY 342 | RA | |
| Oy | 797 LAQLOQVEPNSPNKKTENFSKTTLGMILPLVNLYCPLGMYTRVCEGST---NLGMP 852 | RA | |
| Db | 343 LSQLOVQSDIAHRYTDLGLKITYTGITLIPMNVTGLFAMVNRIGESTDENGPNF3MF 402 | RA | |
| Oy | 853 FGIVELPLITGTSPTIAQWNLKYLANSITCQNNNGRPFIENSSRSIRSISLGLK 907 | RA | |
| Db | 403 FGILGYLCLITVATUAKWLNAYKEKIRGN-----FRNSASDNTSRSMSKR 451 | RA | |
| RESULT 1 | 04 PHQ2_USTMA PRELIMINARY; PRT: 881 AA. | RA | |
| ID | OAPHQ2_USTMA | | |

| | | |
|-------------|--|--|
| Oy | 475 -----TWLDCTCPDSEERKLAKAFIHPULATEDIRMQST 510 | 54 VVKSEKOOHHQEITSDNAKPLTRKSSS1KKKNLTDKDRITNPMSLSGGDTTINSGHK 113 |
| Db | 549 GISATRALKALPGGTRMDPDEANTWMLDVSPDQEMKLJLSVFIHPULATEDILMEST 608 | 63 LLOQVOVDALE-----AGFSIDFQAI-----ADDRSTK----- |
| Oy | 511 REKVLPKSYTFVFHPEAKEDSELEPUNYTVPHDFGILTTPHSPTSHPANVRRV 570 | 114 NRNNYMSLRLKDYL-----KDNDDNSTNNNHTHLAIIPIPIPTITNNKNSRKRKSOLENL 171 |
| Db | 609 RKEIEFRAYHUVCRFSRSDQDYEPTPLEPANNITYFRETSUFSIFRGTPFQNVRAR 668 | 96 ----DNGLGRFLSDPCTGNRCRAFSRTHQI-----ANMSRDSS----- |
| Oy | 571 RQLRDYDVSADWLYCALDIBITGPAVTIGTEYEADEDATFARDTDSMMIQRIG 630 | 172 PPLIKKTGTTGNNSNPNFENDUSPMTRKNTDSDDITNTSTANHMKLGIGATTUGVG 231 |
| Db | 669 KQLKDYINTVNTSWIYALIDDTAFCPLQSTIEYEDSIDEVLVLLKAEGSDMLRG 728 | 134 ----VSARSTSPNSYEAF-----ADPERERANTE-----SHAAPDIEALTLQQTVCSC-GTH 181 |
| Oy | 631 ESRKYMUTMLRLLSKKADVIMCPAKRCOEZANSSGYYQRQYNLUQQQQQAPPNPNTI 690 | 232 TTATATAAAAGR-RPSRSSIDEADSHASRSOFTEDVEDCVPMMVGDHRVNGIDPDETD 290 |
| Db | 729 TCKKVMGLRLMGRADUAVGAKRNCNEWS-----760 | 182 FRRPFNSAFAIRFOFGDQDS-----PEDCATPFEQLGRIVDTEEL 228 |
| Oy | 691 TSPINSTLNLSLGTSTGGVGGINGPNTNTTGSPSPQQQQORGITN 750 | 291 EPRIEREEAYLQKOMIAKNLBDEFQNSKUNNTSGASRHYYHHHSNNNNKANGDGG 350 |
| Db | 761 -----760 | 229 EFV-----ALQKTKPLTTRKHELSQSKK-----254 |
| Oy | 751 KSPPIPDARRADIALYLGQDQHLMFQNLLAYEKIERSHSNSNLAQLOQESNSNPK 810 | 351 GSSMAALKYTPKNIKKTLSFEEFHENSSEBEVLTKQOPPKYDQLSLTSSTPS 410 |
| Db | 761 -----VAPSDGTLIYSDQDHLTMQNLHAYEKILSRNSNLAQISTEMTDAHQ 813 | 255 -----SVFVTDQRPNAQKSAHBAKRSSSEVSESSDDIGLDKDB 295 |
| Oy | 811 ITEMFSKTTLGMPLUNLVYGLGRHNVVRPGEGTNLGWFGTIVGVLIFTIGSFPA 870 | 411 TSGSGSGCVKFGARIISDGINGSLPDRFLSFHSESEETHADPISLVSPLCQSTRDLPLR 470 |
| Db | 814 INDVLSKLTALSTVIVPNTVIGLGRHNVVPGECVENLRWAGTCVSMVAVSYTA 873 | 296 KVFDAYNEK----OLYENURNEMEPREGFPPSSOSVTHAELCDLIVPEPDRDLFQ 351 |
| Oy | 871 QWALKK 876 | 421 NGEE-TWLLDQCTPOSEMKAKGAPIPLATEDTRQESTREKBLKSTPYPCHP 528 |
| Db | 874 TRYLER 879 | 352 LGPEGQWWLWDLNPAITEAEY/GALSRAFSIPLTEDILTBAREKYLPLKQTPVCPRTF 411 |
| Oy | RESULT 15 | 529 -BADKESEDYLEPINVTVVPHDGLTIPHFSPIHSPANTVRYRVRQDUDYSDADMLYCA 587 |
| ID | Q2U77 ASPOR PRELIMINARY; PRT: 756 AA. | 412 YDQDKTSRSEHEPNPNTMIVPRQOLSPSTENPHASNYTRRIGKRDYVLSMDICYA 471 |
| AC | Q2U77_1 ASPOR | 588 LIDETTOQFAPVINGIEVADAEAVPTARDPFSMMLQRTGESRKRKNTLMRSLGKA 647 |
| DT | 24-JAN-2006, integrated into UniProtKB/TREMBL. | 472 MIDDIVDSFGVIRBIEVESEATEDLYFARMDDFSPLPRIGLAKKYSMSLRIGGKA 531 |
| DT | 24-JAN-2006, sequence version 1. | 648 DVIKMFARCKRCOEANSSSGYVQRQYNLQQQQQAPPNNPNTSPINSTLNLSGFTST 707 |
| DE | Mg2+ and Co2+ transporters. | 532 DVTRGSPSKRCNB-----543 |
| CN | ORPName=A0909009006681; | |
| CC | Eukaryota; Fungi; Ascomycota; Pezizomycetes; Eurotiomycetes; OC Eurotiales; Trichocomaceae; mitosporic trichocomaceae; Aspergillus. | 708 GGGVGVGINFQNPNTNTGSPPOOOQGIRTKSPIDARPADIALY 767 |
| OX | NCBI_TaxID:5062; | 544 -----QSVT-----PRBDIGV 556 |
| RN | (1) | 688 LGDIODHITTMFONLLAYEKFSRSHSNLYAQLOVESSENNSNKITEMPSKTKLIGMLVP 827 |
| RP | NUCLEOTIDE SEQUENCE. | 557 LGDIODHITTMFONLLAYEKFSRSHSNLYAQLOVESSENNSNKITEMPSKTKLIGMLVP 616 |
| RA | STRAIN=IB_40; | |
| RA | DOI=10.1038/nature04300; | |
| RA | PubMed=16372010; | |
| RA | Nachida M., Asai K., Sano M., Tanaka T., Kumagai T., Terai G., | |
| RA | Kusumoto K., Arima T., Akita O., Kashiwagi Y., Abe K., Gomi K., | |
| RA | Horiuchi H., Kitamoto K., Kobayashi T., Takeuchi M., Denning D.W., | |
| RA | Galagan J.E., Nieman R., Archer J., J. Bennett J.W., | |
| RA | Bhatnagar D., Cleveland T.E., Fedorova N.D., Gotoh O., Horikawa H., | |
| RA | Hosoyama A., Ichinomiya M., Igarashi R., Iwashita K., Juvvadi P.R., | |
| RA | Kato M., Kato Y., Kin T., Kokubun A., Maeda H., Maseyama N., | |
| RA | Maruyama J., Nagasaki T., Nakajima T., Oda K., Okada K., Paulsen I., | |
| RA | Sawano T., Takemoto T., Takase K., Terabayashi Y., | |
| RA | Wortman J.R., Yamada O., Yamagata H., Anazawa H., Hata Y., Koide Y., | |
| RA | Konori T., Koyama Y., Mineo T., Suharnan S., Tanaka A., Isono K., | |
| RA | Kuhara S., Ogasawara N., Kikuchi H., | |
| RA | *Genome sequencing and analysis of Aspergillus oryzae.*; | |
| RA | Nature 438:1157-1161(2005). | |
| CC | Copyright by the UniProt Consortium, see http://www.uniprot.org/terms | |
| CC | Distributed under the Creative Commons Attribution-NoDerivs license | |
| DR | EMBL: A007150; BAES5038..1; Genomic_DNA. | |
| SEQUENCE | 756 AA; 85231 MW; 864BG2C7D3JB528 CRC64; | |
| Query Match | 20 58; Score 990; DB 2; Length 756; | |
| CC | Basic Local Similarity 31.2%; Pid: No. 8.7e-68; | |
| CC | Matches 265; Conservative 104; Mismatches 234; Indels 246; Gaps 21; | |
| Oy | 4 SESYQNSTTQIPRSDEVIDIRNQI-----TMDCAI-----DSEDBLELKSELSE 53 | |
| Db | 7 SEKRNFLSRSTPVPE----LDQHFRQFLDPPIREATPDMLTSRONTPLNMEPOSD 62 | |

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biacceleration Ltd.

Om protein - protein search, using sw model

Run on: June 13, 2006, 14:05:02 ; Search time 47 Seconds
(without alignments)
1887.486 Million cell updates/sec

Title: US-10-018-105A-4
Perfect score: 4820

Sequence: 1 MSDSEYYQNNTTQPIPRS.....LGKXKGGNKTSISFPNPKY 922

Scoring table: BLOSUM62
Gapext 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 10%
Listing first 45 summaries

Database : PIR_80.*
1: Pir1;
2: Pir2;
3: Pir3;
4: Pir4;

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--|
| 1 | 1680 | 34.9 | 859 | 2 S66827 | aluminum resistance protein AIL1 - yeast (Saccharomyces cerevisiae) |
| 2 | 1666 | 34.6 | 858 | 2 S66825 | N; Alternative names: protein 00522; protein YOL130W |
| 3 | 1173 | 24.3 | 803 | 2 T4036 | C; Species: Saccharomyces cerevisiae |
| 4 | 949 | 19.7 | 451 | 2 S71983 | C; Date: 12-Jul-1996 #Sequence Revision 12-Jul-1996 #text_change 09-Jul-2004 |
| 5 | 865 | 17.9 | 969 | 2 S77886 | C; Accession: S66827; S63340; S71978 |
| 6 | 298 | 6.2 | 387 | 2 S77981 | R; Arino, J.; Casamayor, A.; Gamo, P.J.; Gancedo, C.; Lafuente, M.J.; Aldea, M.; Casas, C.; Herrero, E.; Arino, J. |
| 7 | 251 | 5.2 | 351 | 2 H72360 | Submitted to the Protein Sequence Database, July 1996 |
| 8 | 223 | 4.6 | 319 | 2 H69805 | A; Reference number: S66814 |
| 9 | 214 | 4.4 | 180 | 2 S71978 | A; Accession: S63440 |
| 10 | 212.5 | 4.4 | 354 | 2 S71978 | A; Molecule type: DNA |
| 11 | 207 | 4.3 | 380 | 2 AG2319 | A; Residues: 1-12, 'Y' 14-859 <CAS> |
| 12 | 200 | 4.1 | 351 | 2 S69348 | A; Experimental source: strain PY1679 |
| 13 | 192 | 4.0 | 1043 | 2 T13733 | A; Cross-references: UNIPARC:UP10000168A52; EMBL:U41293; NID:gi1209710; |
| 14 | 186.5 | 3.9 | 1072 | 2 A86827 | R; Khalid, H.; Balcels, L.; Aldea, M.; Casas, C.; Herrero, E.; Arino, J. |
| 15 | 184.5 | 3.8 | 1138 | 2 T18287 | Submitted to the EMBL Data Library, November 1995 |
| 16 | 180 | 3.7 | 1212 | 2 T1804 | A; Description: Sequencing of a 13.4 kbp fragment of the left arm of chromosome XV reveals a putative ser/thr protein kinase gene, the malate dehydrogenase MDH2 gene, a ribosomal L25 gene. |
| 17 | 178.5 | 3.7 | 375 | 2 DT591 | A; Reference number: S63440 |
| 18 | 177.5 | 3.7 | 2452 | 1 RN202L | A; Status: nucleic acid sequence not shown; translation not shown |
| 19 | 177 | 3.7 | 2457 | 2 T18492 | A; Molecule type: DNA |
| 20 | 176 | 3.7 | 1585 | 2 T18274 | A; Residues: 1-12, 'Y' 14-859 <CAW> |
| 21 | 172 | 3.6 | 1091 | 2 S57112 | A; Cross-references: UNIPARC:UP10000168A52; EMBL:U41293; NID:gi1209710; |
| 22 | 171 | 3.5 | 1368 | 2 T18371 | PIDN: AAC94622.1; PID:gi1209711 |
| 23 | 170 | 3.5 | 1416 | 2 D71639 | A; Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995 |
| 24 | 169 | 3.5 | 322 | 2 C87643 | C; Genetics: |
| 25 | 166.5 | 3.5 | 1858 | 2 T18273 | A; Gene: SGD:ALR1; ALR1 |
| 26 | 165.5 | 3.4 | 1127 | 2 T28317 | A; Cross-references: SGD:S00005490 |
| 27 | 164.5 | 3.4 | 2271 | 2 F9073 | A; Map Position: 15L |
| 28 | 163.5 | 3.4 | 956 | 2 S30834 | C; Keywords: transmembrane protein P743-759; Domain: transmembrane #status predicted F777-793/Domain: transmembrane #status predicted C; Genetics: |
| 29 | 163 | 3.4 | 852 | 2 T06310 | Query Match 34.9%; Score 1680; DB 2; Length 859; |
| 30 | 162 | 3.4 | 1570 | 2 T18222 | Best Local Similarity 41.5%; Pred. No. 9-95; Mismatches 258; Indels 176; Gaps 31; |
| 31 | 161 | 3.3 | 366 | 2 F70932 | A; Cross-references: SGD:S00005490 |
| 32 | 160 | 3.3 | 369 | 2 D87045 | C; Map Position: 13L |
| 33 | 160 | 3.3 | 373 | 2 T45262 | Query Match 58%; Score 1680; DB 2; Length 859; |
| 34 | 159.5 | 3.3 | 1584 | 2 T18276 | Best Local Similarity 41.5%; Pred. No. 9-95; Mismatches 258; Indels 176; Gaps 31; |
| 35 | 159.5 | 3.3 | 2150 | 2 S71629 | A; Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995 |

C:Species: *Saccharomyces cerevisiae*
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004
C:Accession: S37886
R:Rasmussen, S.; von Wettstein, D.
submitted to the Protein Sequence Database, March 1994
A:Reference number: S37886
A:Accession: S37886
A:Molecule type: DNA
A:Cross references: UNIPROT:P35724; UNIPARC:UPI000012F2A4; EMBL:Z228064; NID:9486086;
PDB:CARB1501_1; PID:9486087; MIPS:YKL064w
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:MRK2
A:Cross references: SGD:S0001547; MIPS:YKL064w
A:Map position: 11L

Query Match 17.9%; Score 865; DB 2; Length 969;
Best Local Similarity 24.2%; Pred. No. 7e-15; Gaps 29;
Matches 271; Conservative 143; N mismatches 247; Indels 460; Gaps 29;

RESULT 6
S75981
manganese/cobalt transport protein s110507 - *Synechocystis* sp. (strain PCC 6803)

Qy 11 STINOPTRSEVLDHHNQNTDCATSDSDELELKSELESEWKSE-----58
Db 44 SDSSRPTQOLLHDLQHNGQIDPQI-DSGMHHESDPSNDLKSAGTDIR 102

Qy 59 ---KQOQHHQBTS-NAKPTUTRSGSIKKNLNTDKDRITNPMS-LSGGDTINTSGH 112
Db 103 PSNSQREGQPSVSTVTPQMFSTSYKPKAGLRSQDPSNSLWSDLSPSB---AEWSNL 160

Qy 113 KNRNTMMASLRDPLYKLN-TDIDSTNNTHLAPITPIPIPTITA-----NKSRIK 165
Db 161 KRP---KSVHHSYDENSPTRQSNANNDVVD-ALMNHRNNNASTGVNDNSKRK 215

Qy 166 SOLENPPPLIKKTKIGRNNSNFFNDLVSPMTKMTNDSDINTTSTIANHMKLGATT 225
Db 216 -----KRP---GDDSSN-KN---SKSTSSDSNDEDEYS-----244

Qy 226 LGVGTCTATATAAGRSRSSTIDEADSHASRSQTEEDCPFM-YGDIHRVG 283
Db 245 -----RSS-----LSENNSLSDDVLLDVLBEGSEYPRAW 275

Qy 284 IDPDEIDPTEBEAYLQKOMIAKNLRIDEFQNLSKNNTGASRIPYHHSNNKK 343
Db 313 D----GTSNEGILPSKIVNIDVPLGNRNRNETENLKGRLPKIAWHJQREM 367

Qy 404 LATSSTSSTSGGCGCGVKGGARISDGINGSSL--PD--RFSLPHES 446
Db 368 VL-----GSNSTDSKSRISQSLQDNLLVGRNQYYPHIISNPHEFRTYRDL 418

Qy 447 EFTIAPDIPSLVSGQVRLF-----469
Db 419 DSTVHSPTISGILQCGQFOFLFATSYODNSAIGHKTHPNSPTCIGAKTVSQQLQT 478

Qy 470 -----RGEET--WWLDCTCPDSEKMLAKAFGTHPLTAEDIIRNQETREKVE 515
Db 479 AKNPSTLSMAYANTEDPFWFLVSNPTEEENKTSKAFGIPLTEDIFGEYREKE 538

Qy 516 LPESTYFCPHTP-----AKDES-----534
Db 539 LFRDYLICPFSFDIVAEKHVRERKEKOBSATDHESTSRRKSOYATMSNEANN 598

Qy 535 -----534
Db 599 NSTNSAERSKWLPSLILARRSISANRTNTSSYYKRYVSEKRRHEENEKPKRSGRH 658

Qy 535 --EDYLPIVNYVPHDGLTHFSPISHPANVERRQRLDYDVSADMLCYALDE 591
Db 659 KRECELPPLAVNVYIIVPRTSVLTHAFTPHPINVRARLLKDYLAVNTSDIAVALDD 718

Qy 592 ITDGAPEHKIEYAAEIAIDAVTARDTDESS-----624
Db 719 ITDAPAPIELEDDEVTEBIALKMHQSDDSDSOSDSDSAGSDEDAPPDVYSK 778

Qy 625 -----MLQIGESEKRWKMTMLRLLSGKADYI 650

Db 779 TSYSSAKSSVSSRSRMSTSEA5FANL17GWRKKGDMRLRIGECKRKVMSIRLGLSGADYI 838
Qy 651 KMPFKRCQBEANSSSGYYQRYN1QCCQQQQA PPPPPNITSPNINSTLNLSUGTGTGG 710
Db 839 KMPFKR-----YNEQ-----849

Qy 711 VCGVGGINFNPNTNTGSPSPQQQHQGITNKSPPIDPARRADIALYLD 770
Db 850 -----EASPOSELAMYLD 863

Qy 771 IODDHITMFQNLAYEKFISHSHENYLAQLOVEFENNNKITEKSKTILGTMVPLNL 830
Db 864 IQDHLYTMVSENLHYEKUHLSSHENYLAQINIDTKVNNDMDVLSKTKTLETLVPMNV 923

Qy 831 VTCLEGMANTRPGEGCTNLGFPGIVGULPIIGSIFAQ 871
Db 924 IYGLWGANVTVPGQRTDSLWFIGVLFMCMASATYMTK 964

A:Variety: PCC 6803
N:Alternative names: protein s110507
C:Species: *Synechocystis* sp.

C:Accession: S75981
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; Hirose, M.; Sugiyama, M.; Sasamoto, S.; Kimura, T.; Matsuno, A.; Tabata, S.; Nakazaki, A.; Naruo, K.; Okumura, S.; Shimojo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, M.; Tabata, S.

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions.

A:Reference number: S74322; PMID:9706201; PMID:8905231

A:Status: nucleic acid sequence not shown; translation not shown

A: Molecule type: DNA
A: Residues: 1-367 <KAN>
A: Cross-references: UNIPROT:Q55481; UNIPARC:UP0000003434; EMBL:D64006; GB:AB001319;
NID:G100129; PID:BA10028.1; PID:G1001341

NID:G100129; PID:BA10028.1; PMID:9706201; PMID:9706201; PMID:8905231

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Superfamily: diivalent cation transporter, Caa type

Query Match 6.2%; Length 387;
Best Local Similarity 23.0%; Pred. No. 8e-11; Gaps 148; Gaps 18;
Matches 108; Conservative 70; Mismatches 143; Indels 148;

Qy 437 DRFLSFH-----SESEETHTADIPS-----LV--SPQSVR--DLFRNGBTW-----W 477
Db 27 DTFDFVDEPSESEGTSIIPDPSRIVLVDSPSHAVRKSDISNALPVLGNTVSW 86

Qy 478 LDCT-CPDSEMVKLAKAFGHPUTADIRMQTRKEVLFKSYTVCFHTEAKDKEBD 536
Db 87 MDIECLGSEEVYLKEVGFPLKPLLEDIVNVPORAKVEDYDNDYVTAHVRPIRE-ED 145

Qy 537 YLEPINV-----YIVVENDGLTIPHPSIISPISPAVNRVRYVQLRDYV-DVSADKUVAL 588
Db 146 GFESEQVSPVQKRYLTFDGH1DOPNPL-----RERITNGDKY/CQGADYLCYL 199

Qy 549 IDETDGSPAVTGHQTEAYDAEDVAYTARDTIDVYRMLLISGAD 648
Db 200 IDMLIDEPYLEDYERIEEALDSTII--RNPN-SSLMEVTHIRRELLARLLI-----251

Qy 649 VIKMFARCKQEANSSGTYQRYN1QCCQQQAPPPIPNTSPINSTLNLSUGTGS 708
Db 252 -----WPLRHMVNLURD-----265

Qy 709 GGVGVGINFNPNTNTGSPSPQQQHQGITNKSPPIDPARRADIALYLD 768
Db 266 -----TNSIVI-----278

Qy 769 GDIQDHITMFQNLAYEKFISHSHENYLAQLOVEFENNNKITEKSKTILGTMVPLNL 828
Db 279 RDCDHTIQVLDIEAYRELASSLMVYKTAM-----SNCONVMKPLTISTIFPL 331

Qy 829 NLVYLGFMNVRVGEGGTNLGFPGIVGVLPIIGSIFAQ---WL 874

| | |
|---|--|
| A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocysis sp. FCC603. II: Sequence determination of the entire genome and assignment of potential protein-coding regions. MUID: 8905231 | Qy 423 GARISDGINGGSSLPDRESLPSESEET IHPADPLSVEFGSQYVRL---PRNGEETW 476 |
| A;Accession: S77053 | Db 7 GVKRKKGKPGSSL---IHVGISTSKTKTEHIIIDYDKEVDRVHSDEGLAYRFDNTN 63 |
| A;Status: nucleic acid sequence not shown; translation not shown | Qy 477 WLDCMCPIPDSE--KMLAKAFGIPHLTAEDTRMQETREKVELPKYSYFVCPHTFEADEKE 535 |
| A;Molecule type: DNA | Db 64 WINTGLEDDIDLYEVKVKYFQEVHSLMEDLNTNQRPKIEEFKFKYFTVULKMYP---RE 120 |
| A;Residues: 1-180 <XANS> | Qy 536 DYLEPINYIVVFDIGLILFH---FSPSHPANTARRVRQLRDYV-DVSADWLCY 586 |
| A;Cross-references: UNIPROT:Q55968; UNIPARC:UPI00000D344C; EMBL:D64005; GB:AB001339; | Db 121 DKLVVEQSVTICMENLTIPDEEGKQODDV---GNLDRTRNTKSIRKTGVYLYT 174 |
| NID:91001779; PID:BA10745_1; PID:91006532 | A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996 |
| C;Superfamily: divalent cation transporter, CorA type | Qy 587 ALIDRITDGAPVHGEYEADATEDAVFTARDTDFSSMLQRIGESRRKVMTLMILSSK 646 |
| Query Match 4.4%; Score 214; DB 2; Length 380; | Db 175 ALIDIVNDVYVILEKEDKTFNEESENT-----S 204 |
| Best Local Similarity 19.28%; Pred. No. 1..1e-05; Matches 89; Conservative 72; Mismatches 141; Indels 140; Gaps 14; | Qy 647 ADVIKMPAKRCOEANSSGGYQROTYNQQOQAPPNPENPTI---TSPINSTNLNS 702 |
| Qy 478 LDCTCPDSE-----MGRHLAKAFGIPHLTAEDTRMQETREKVELPKYSYFVRL 522 | Db 205 -----TATNNLFNEVYNLKLQMQ-----SASNKKNEELMKPFTTIPS 301 |
| Db 59 LDEAVDVTESVSWINIDGLGNHNTWEGLEPEVHEVALEDTVNTQPKVYEVNLH 118 | Qy 703 LGTSTGCGGVGGVINFGPNCRNTNTNTGSPSPQQQHQGITNSKSPPTPDRPRA 762 |
| Qy 523 VCPHPEADKSESDYLEDPEINVVYIVVENDGILTFPSP-ISHPANVRVNRVQLRDYV-DVS 580 | Db 237 -TQRGEVET-----S 248 |
| Db 119 FISRATVTLQSSQTTISQFISIFLQKH-YLTIDCEPKYDLSVRSRIRTKGAIRQKN 177 | Qy 763 DIALYLGDOIHDITMFONLILAYERKIFSRSHSYLAQLOVESFSNNKITEMPSKTLIG 822 |
| Qy 581 ADMLCYALIDESTDGFAPVHGEYEADATEDAVFTARDTDFSSMLQRIGESRRKVMTLM 640 | Db 249 DISITPKVQYHTQVLDITTEPLDMTAGLDTYL-----SASNKANEELMKPFTTIPS 301 |
| Db 178 ADLFLYALIDATIDGFFPVV---MEVYGLAQVSLQSEIISCPDKTSLAKTHQDQDLTIR 234 | Qy 823 TMLPLNVLTVGLEMATYRVPQEGTNLGMFPCTGVLGULIIGSFPAQMWLKK 876 |
| Qy 641 RLLSGKADYKRFARCOBEANSSGTYQROYNQQQQAAPPNPNTSPINSTNL 700 | Db 302 TTFBLSPFLVGVGMNFDMPE---LKFRYGTAVLWIMISLAPMAYFRK 351 |
| Db 235 RAJWPORDAI-----244 | RESULT 11 |
| Qy 701 NSLIGSTGCGVGGVINFGPNTGPNTPGNTNTNTGSPSPQQQHQGITNSKSPPTDARP 760 | AG2319 |
| Db 245 NSL-----257 | hypothetical protein al14110 [Imported] - Nostoc sp. (strain PCC 7120) |
| Qy 761 RADIALYLGDOIHDITMFONLILAYERKIFSRSHSYLAQLOVESFSNNKITEMPSKTL 820 | C;Species: Nostoc sp. PCC 7120 |
| Db 251B --EVYTFURDYTDHTIQIQLDTEITYRDYSLSSY-----SRHMEIMKLTIV 308 | A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120 |
| Qy 821 IGMVLPAINLTVGLEMATYRVPQEGTNL--LGMPFG---IVGVLIPIIGTSFTAOW--- 872 | C;Accession: AG2319 |
| Db 309 ISSIPPLTIDIGTGYMFNPKRSPPNMPPELAWKWWGYPTVNMMLTVGMMLYFWRKG 367 | R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriuchi, M.; Ishikawa, A.; Kawashita, K.; Murakami, A.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Matsunoto, M.; Matsumoto, M.; Takazawa, M.; Yamada, M.; Yesuda, M.; Tabata, S. |
| Qy 873 WURKLNNSTEGGNNR 889 | DNA Res. , 005-213, 2001 |
| Db 368 WFPNLND---VEKNR 380 | A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena sp. strain PCC 7120 |
| A;Reference number: AB1807; MUID:21595285; PMID:11759840 | A;Accession: AG2319 |
| A;Status: preliminary | A;Status: preliminary |
| A;Molecule type: DNA | A;Molecule type: DNA |
| A;Residues: 1-380 <KUR> | A;Cross-references: UNIPROT:Q8YPT1; UNIPARC:UPI00000CEA09; GB:BA000019; |
| PIDN:RAB75809_1; PID:1733345; GSDB:GN0179 | C;Accession: Sep-2001 #sequence_revision 14-Sep-2001 #text_change 31-Dec-2004 |
| A;Experimental source: strain PCC 7120 | C;Genetics: |
| A;Gene: al14110 | C;Superfamily: divalent cation transporter, CorA type |
| A;Accession: E97128 | Qy 477 WLDCMCPIPDSE--KMLAKAFGIPHLTAEDTRMQETREKVELPKYSYFVCPHTFEADEKE 534 |
| A;Status: preliminary | Qy 478 WVDQGLGNDQNLDRVSQVTFPLPSVEDDNNSERPKTEDDQILFIAHMVKERES 133 |
| A;Molecule type: DNA | Db 535 EDYLEPINYIVVFDIGLILTHPSIHP---ANVARYRQLRDYV-DVSADMLCYALDE 591 |
| A;Residues: 1-354 <KUR> | Db 134 GFYSEQVS---IILCKHYLLSQEP---EHOCFAGVNRRIKGKOTRQADYLAYALLDA 190 |
| A;Cross-references: UNIPROT:Q97104; UNIPARC:UPI0000CA31D; GB:AE001437; | Qy 532 ITDGPAPVHGEYEADIRD----AVPDPARDDESSMLQRIGESRRKVMTLMILSSK 646 |
| PID:G1524829; GSDB:GN00168 | Db 131 IIDGFPV---LERYGEALELBBEVVSPSQT---LQKTHYTRRLQQRRAINFO 242 |
| A;Experimental source: Clostridium acetobutylicum | Qy 667 ADVIKMPAKRCOEANSSGGYQROYNQQQQAAPPNPNTSPINSTNLNSLGTS 706 |
| C;Genetics: | Db 243 RDATNL-----IRDRSQLISP-----259 |
| A;Gene: CAC1052 | Q;Score 212.5; DB 2; Length 354; |
| C;Superfamily: divalent cation transporter, CorA type | Q;Best Local Similarity 18.8%; Pred. No. 1..3e-05; Matches 77; Mismatches 159; Indels 149; Gaps 15; |
| Q;Score 214; DB 2; Length 380; | Q;Best Local Similarity 18.8%; Pred. No. 1..3e-05; Matches 89; Conservative 77; Mismatches 159; Indels 140; Gaps 14; |

| | | |
|-----|---|-----|
| 552 | ILTPHFSPLSHPANVERVRQLRDYDVSDMIVVALIDEITDGPVTRGIVYEADAE | 611 |
| 541 | --LIDE-----GYENADSE | 554 |
| 612 | DAVPTARDTDFSSMLQRTGESBRKVNTMLRLSGKA- | 656 |
| 555 | --NLSBTKRIN--KTEIENLNTLISNSNSWNSLNGSSTTCINPLSP | 602 |
| 657 | CQBEBANSCTYORQNLQOQQQQAAPPPIPPI-ITSPINSTLNLSLGT--STGGCG | 712 |
| 603 | SMMNPSSSTSSNTLNLSURKFSQLKLRAELYSINTSNSRGASSTSGCSGSG | 662 |
| 713 | VGGCINFPRNPQGNTNTTNT-NTGCSPSPPQQQCGITNKSFPIIDASPRAD | 763 |
| 663 | GGRNRNCFTSSNSNSPIOPFENEDISTBKNEBPFESTLTLI-----RB | 713 |
| 764 | IALYLGDIQHITMEQNULAYKEFKFSRSHNLAQLOV--ESPNMNK | 810 |
| 714 | NOLYKTE-----PKEIKLKELESNSKSSNNIWOLEYKSTQVLKPKDQSDRNKT | 767 |
| 811 | ITEMSKITLIGMTALPMALVTLGLGAVNTPFGGCGTNGPFFGIVVFLIGSFIGA | 870 |
| 768 | RKQLFN-----GSN-----GSN----- | 781 |
| 871 | QWNLKLUKLNIECONNINGRPENHNSRSRSIRSLSGLKJHGKNSKTSIPNPKY | 921 |
| 782 | --NSGSSGCRHHHHHCNN-----NGSMSETI--PSY | 811 |

Search completed: June 13, 2006, 14:09:04
Job time : 52 secs